UNIVERSITY OF TORONTO

Faculty of Arts and Science

August 2018 EXAMINATIONS

STA261H1S

Duration - 3 Hours

Aids Allowed: Non-programmable calculator

First Name:____

Last Name:____

Student Number:____

U of T Email:____

This exam booklet contains 23 pages. Write all final answers in this exam booklet, on the same page as the question appears. Use the 4 pages at the end for rough work. Answer all questions in pen. Do not write anywhere near the QR code at the top of each page, this will affect the scanning of your exam. Aids permitted: non-programmable calculator.

Questions:

Question	Marks Possible
1	16
2	20
3	24
4	20
5	24
Total	104

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 θ and θ_0 is the true value then

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

and

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

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,

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

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

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

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

$$\frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

If $W_n \sim \chi^2_n$ then $E(W_n) = n$ and $Var(W_n) = 2n$.

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

1. (16 marks) Let X be a random variable taking values on the whole real line, with density depending on parameter $\theta \in \Omega \subset \mathbb{R}$ given by

$$f(x; \theta) = \exp\left(\theta x - b(\theta)\right) h(x)$$

where $h(x)$ is a function depending on x but not θ and $b(\theta)$ is a function depending on θ but not x.

(a) (4 marks) Let $x_i, i = 1 \ldots n$ be an IID random sample from this distribution. Find the log-likelihood and the score statistic for θ.

(b) (2 marks) Find a sufficient statistic for θ

(c) (6 marks) Show that $E(X) = b'(\theta)$ (where the derivative is with respect to θ). Hint:

$$
\begin{aligned}
&\int_{-\infty}^{\infty} f(x;\theta)dx = 1 \\
\implies &\int_{-\infty}^{\infty} \exp(x\theta)\, h(x)dx = \exp\big(b(\theta)\big)
\end{aligned}
$$

You may assume any necessary mathematical conditions required to exchange the order of differentiation and integration.

(d) (4 marks) Find the Method of Moments estimator of θ, and show that it equals the Maximum Likelihood Estimator. You may assume that $b'(\theta)$ has a unique inverse in Ω.

2. (20 marks) The amount of rainfall in inches was recorded for 227 storms in Illinois from 1960 - 1964. We wish to fit a probability model to these data. We have two candidates in mind, a $Gamma(\alpha, \lambda)$ distribution and a simpler $Exponential(\lambda)$ distribution. Recall $Gamma(1, \lambda) \stackrel{d}{=} Exponential(\lambda)$, with pdfs as follows:

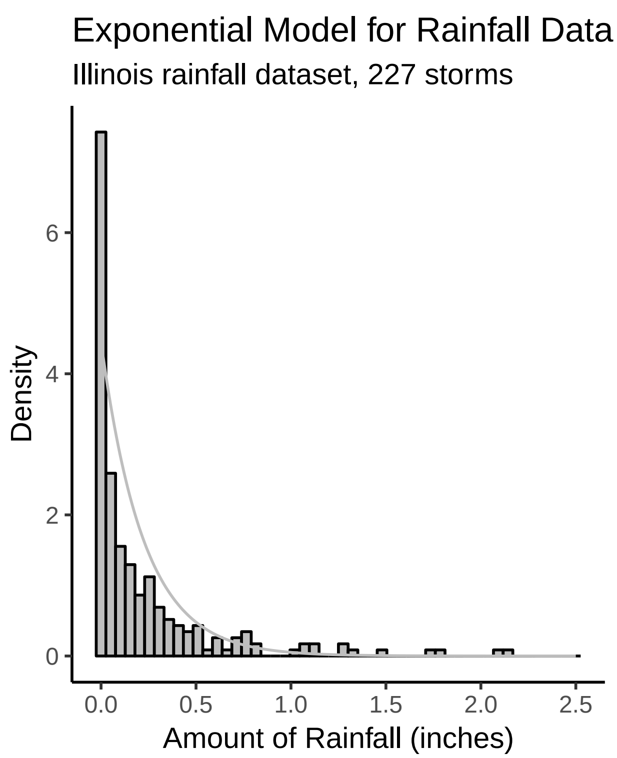
$$\text{Gamma: } f(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x)$$

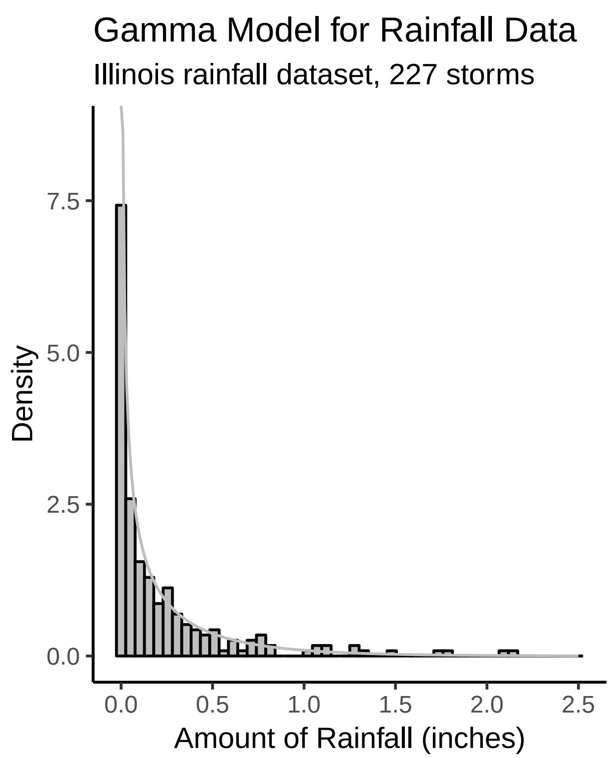
$$\text{Exponential: } f(x; \lambda) = \lambda \exp(-\lambda x)$$

We are interested in whether the exponential model fits the data well enough, or whether the gamma model is necessary.

We fit both curves by maximum likelihood, obtaining the result shown in the histograms entitled "Exponential Model for Rainfall Data" and "Gamma Model for Rainfall Data".

```
## Exponential distribution, MLE for lambda =  4.456485

## Gamma distribution, MLE for alpha =  0.4407903

## Gamma distribution, MLE for lambda =  1.964367

## Rainfall data, sum of x =  50.937

## Rainfall data, sum of log(x) =  -672.87

## Gamma(.441) =  2.008587
```





(20 marks) Perform a Likelihood Ratio Test to investigate whether the Exponential model is appropriate for these data. Show all steps and make a full conclusion in plain, non-technical language.

(Question 2 continued)

3. (24 marks) Let $X_i \overset{IID}{\sim} N(\mu, \sigma^2)$. We wish to estimate the variance $Var(X) = \sigma^2$. We consider two estimators:

$$s_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

$$s_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

We saw before that s_n^2 is the maximum likelihood estimator of σ^2. The estimator s_{n-1}^2 is often taught in statistics courses as being "correct", because it "corrects bias" in s_n^2. This being a math stats course, let's go a bit further and compare the properties of these estimators.

(a) (4 marks) Compute the bias of s_n^2 and s_{n-1}^2. Is either estimator unbiased?

(b) (4 marks) Compute the variance of s_n^2 and s_{n-1}^2. Is one always higher than the other, or is it impossible to say?

(c) (6 marks) Find the Fisher information $I(\sigma^2)$ in the sample. Is either estimator efficient?

(d) (4 marks) Compute the Mean Squared Error of both estimators. Is one always lower than the other, or can you not say?

(e) (6 marks) Discuss the relative merits of each estimator. Give at least one positive and one negative aspect of each. You will be marked on the clarity and thoroughness of your discussion.

4. (20 marks) Data on the annual temperature measured in Ann Arbor, Michigan, is shown in the plots entitled "Histogram of Annual Temperature" and "Temperature by Year". Overlayed on the histogram is a Normal density curve; the normal model seems to fit quite well. We know how to estimate the mean μ and standard deviation σ of this distribution. However, scientists want to know: is the average temperature increasing across years?

You propose the model

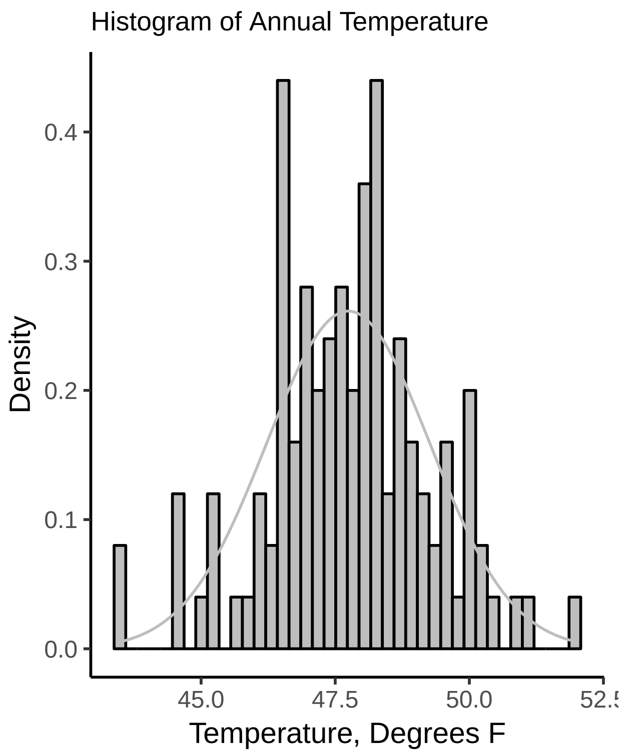
$$Y_i \overset{ind}{\sim} N(\mu_i, \sigma^2)$$
$$\mu_i = \alpha + \beta x_i$$

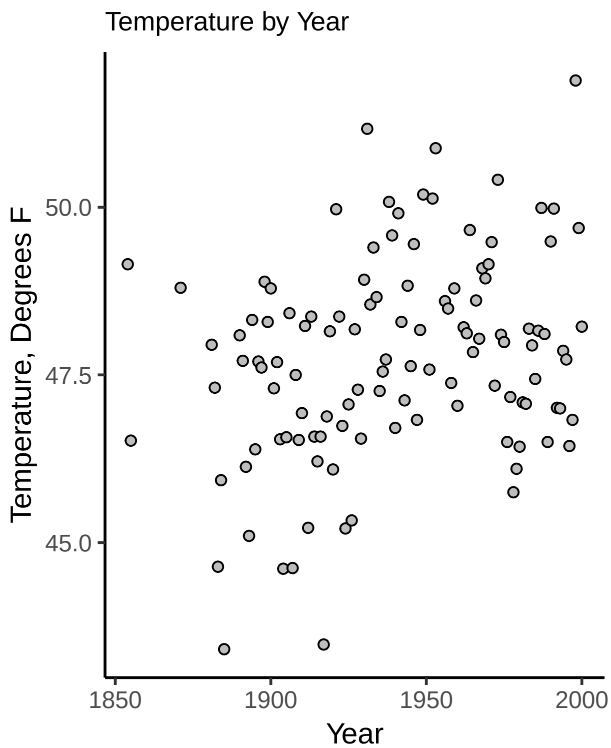
where Y_i is the temperature in the ith year, with year 1 being 1854; $x_i = 1, 2, \ldots 115$ is the i^{th} year (i.e. $x_1 = 1854, x_{115} = 2000$), and α, β, σ^2 are parameters to be estimated. Under this model, it is clear that β has the interpretation as the expected increase in temperature in any pair of consecutive years, since

$$\mu_{i+1} - \mu_i = \beta$$

Hence if we can estimate β, we can tell the scientists whether their hypothesis is reasonable.

```
## Warning: package 'bindrcpp' was built under R version 3.4.4
```





(a) (2 marks) Write down the log-likelihood for α, β. You can treat σ^2 as a fixed, known constant for this question.

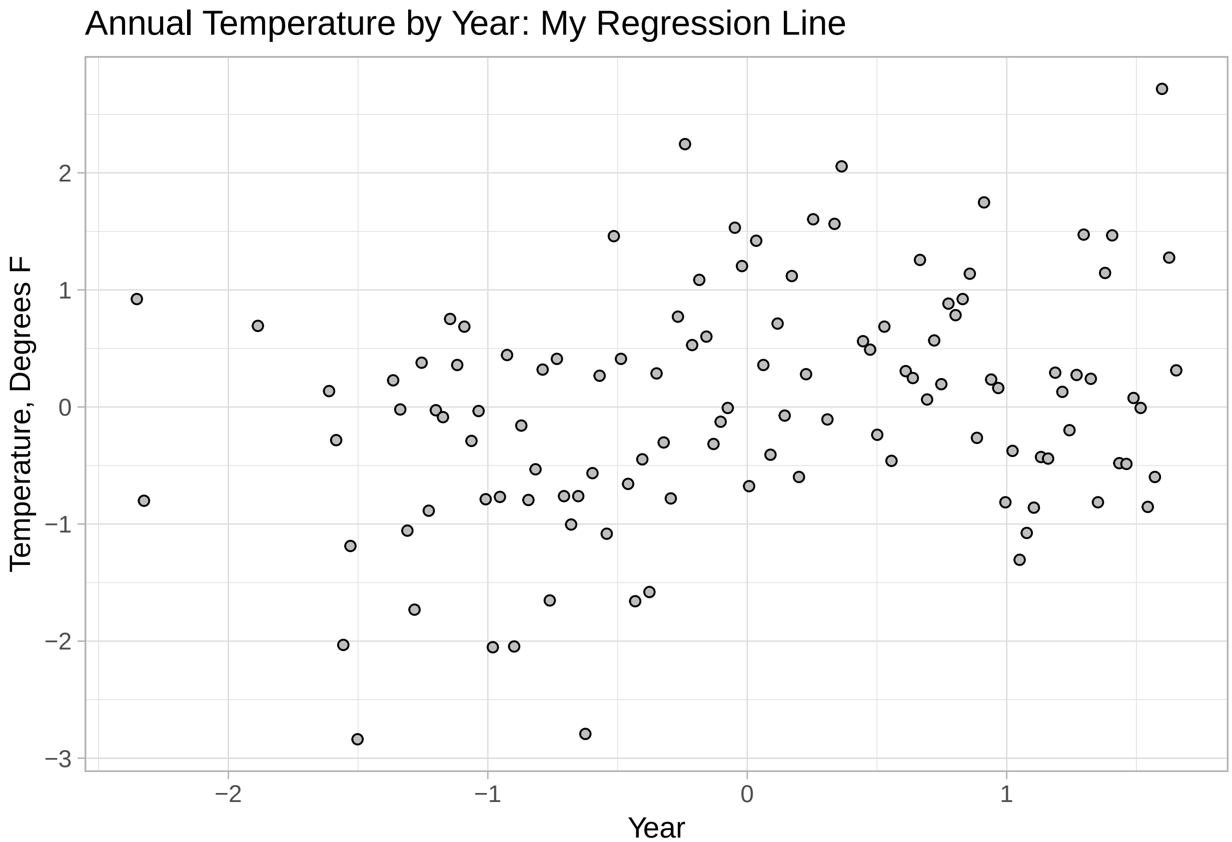
(b) (2 marks) Find the score statistics for α, β.

(c) (8 marks) Find the Maximum Likelihood Estimators for α, β, and compute them for these data. I have centred and scaled the data, so you may use

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = 0$$

$$\sum_{i=1}^{n} x_i^2 = 114; \sum_{i=1}^{n} x_i y_i = 33.31$$

(a) (8 marks) Consider the plot entitled "Annual Temperature in Ann Arbor by Year: My Regression Line". You just fit a linear regression- modelling $E(Y_i|x_i) = \alpha + \beta x_i$. Draw a line on the plot with the slope and intercept you found in the previous steps. The data in the plot has been centred and scaled as described in the previous part.



5. (24 marks) Let $X_i \sim Exp(\theta), i = 1 \dots n$ be an IID random sample from an exponential distribution with the parametrization

$$f(x|\theta_0) = \theta_0 e^{-x\theta_0}$$

where θ_0 is the true value of θ.

(a) (8 marks) Find $E(X)$, $Var(X)$, the MLE, and the Observed Information.

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

(c) (3 marks) Use the Central Limit Theorem for the sample mean to find an approximate 95% confidence interval for $1/\theta$.

(d) (3 marks) Use your answer to the previous question to find an approximate 95% confidence interval for θ. Call this interval V_n

(e) (3 marks) Use the Central Limit Theorem for the MLE to find an approximate 95% confidence interval for θ. Call this interval W_n.

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

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

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

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

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