Problem 1

Let $X_1, \dots, X_n \overset{i.i.d.}{\sim} N\left(\mu, \sigma^2\right)$. We will use the generalized likilihood ratio method to construct a test for the hypotheses:

$$H_0 : \mu \leqslant 5 \quad \text{vs.} \quad H_1 : \mu > 5$$

The likilihood function is given by:

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

The log-likilihood function is therefore:

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

Computing the MLEs $\hat{\mu}, \hat{\sigma}^2$ gives:

$$\frac{\partial LL}{\partial \mu} = 0 \Rightarrow \sum_{i=1}^{n} x_i - n\mu = 0 \Rightarrow \hat{\mu} = \bar{X}_n$$

$$\frac{\partial LL}{\partial \sigma^2} = 0 \Rightarrow -\frac{n}{2\sigma^2} - \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2$$

The MLEs $\tilde{\mu}, \tilde{\sigma}^2$ constrained by the null hypothesis will be identical to $\hat{\mu}, \hat{\sigma}^2$ if $\bar{X}_n \le 5$; and $\tilde{\mu} = 5, \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-5)^2$ otherwise. If the two sets of MLEs are identical, then $\Lambda = 1$, and we would never reject H_0. Therefore, to determine a decision rule for rejection, we may just consider the case where $\tilde{\mu} = 5, \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-5)^2$. From our MLEs, Λ is given by:

$$\Lambda = \frac{L\left(\tilde{\mu}, \tilde{\sigma}^2\right)}{L\left(\hat{\mu}, \hat{\sigma}^2\right)} = \frac{\left(2\pi\tilde{\sigma}^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\tilde{\sigma}^2}\sum_{i=1}^{n}(x_i-5)^2\right\}}{\left(2\pi\hat{\sigma}^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2\right\}} = \left(\frac{2\pi\hat{\sigma}^2}{2\pi\tilde{\sigma}^2}\right)^{\frac{n}{2}}$$

$$= \left(\frac{2\pi\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2}{2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-5)^2}\right)^{\frac{n}{2}} = \left(\frac{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2}{\sum_{i=1}^{n}(x_i-5)^2}\right)^{\frac{n}{2}}$$

Our decision rule is to reject H_0 if and only if $\Lambda < c \le 1$ where is is determined by $P_{\mu\leqslant 5}(\Lambda < c) = \alpha$. Manipulation of the test statistic and decision rule gives:

$$\left(\frac{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2}{\sum_{i=1}^{n}(x_i-5)^2}\right)^{\frac{n}{2}} < c \Rightarrow \left(\frac{\sum_{i=1}^{n}(x_i-5)^2}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2}\right)^{\frac{n}{2}} > c' \Rightarrow \frac{\sum_{i=1}^{n}\left(x_i-\bar{X}_n+\bar{X}_n-5\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} \geqslant c''$$

$$\Rightarrow \frac{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2 + 2\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)\left(\bar{X}_n-5\right) + \sum_{i=1}^{n}\left(\bar{X}_n-5\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} > c'''$$

$$\Rightarrow \frac{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2 + n\left(\bar{X}_n-5\right)}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} > c''' \Rightarrow 1 + \frac{n\left(\bar{X}_n-5\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} > c''''$$

$$\Rightarrow \frac{\left(\bar{X}_n-5\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2/n}} > c''''' \Rightarrow \frac{u}{\sqrt{V/d}} \sim T(n-1)$$

So, the simplified test statistic $\frac{\sqrt{n}\left(\bar{X}_n-5\right)}{S}$ is recognized as having a Student's T distribution with $n-1$ degrees of freedom. Therefore, the test function may be written as:

$$\phi(Z_1, \dots, Z_n) = \begin{cases} 1 & \text{if } \bar{X}_n > 5 + \frac{C_{0.05,n-1}S}{\sqrt{n}} \\ 0 & o.w. \end{cases}$$

where $C_{0.05,n-1}$ is determined from the Student's T distribtion with $n-1$ degrees of freedom such that the area under the density to the right of $t = C_{0.05,n-1}$ has an area of $0.05 = \alpha$.

Problem 2

Let $Y_i = \alpha + \beta x_i + \varepsilon_i, i = 1, 2, \dots, n$, with $\varepsilon_1, \dots, \varepsilon_n \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$ and σ^2 is unknown. Assume $n \ge 3$. It is required to find a generalized likelihood ratio test for:

$$H : \beta \leqslant \beta_0 \quad \text{vs.} \quad K : \beta > \beta_0$$

Since $Y_i \sim N\left(\alpha + \beta x_i, \sigma^2\right)$, the likelihood function is given by:

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - (\alpha+\beta x_i)\right)^2\right\}$$

$$= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha+\beta x_i)\right]^2\right\}$$

So, the log-likelihood function is given as:

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha+\beta x_i)\right]^2$$

The MLEs for α, β, σ^2 are found by simultaneously solving:

$$\begin{cases} \frac{\partial \ln(L)}{\partial \alpha} = 0 \\ \frac{\partial \ln(L)}{\partial \beta} = 0 \\ \frac{\partial \ln(L)}{\partial \sigma^2} = 0 \end{cases} \Rightarrow \begin{cases} \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha+\beta x_i)\right] = 0 \\ \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i\left[y_i - (\alpha+\beta x_i)\right] = 0 \\ \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha+\beta x_i)\right]^2 - \frac{n}{2\sigma^2} = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{1}{n}n\alpha - \frac{\beta}{n}\sum_{i=1}^{n} x_i = 0 \\ \frac{\sum_{i=1}^{n} y_i\left(x_i-\bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} - \beta = 0 \\ \frac{1}{n}\sum_{i=1}^{n}\left[y_i - (\alpha+\beta x_i)\right]^2 - \sigma^2 = 0 \end{cases} \Rightarrow \begin{cases} \hat{\alpha} = \bar{Y}_n - \hat{\beta}\bar{X}_n \\ \hat{\beta} = \frac{\sum_{i=1}^{n}\left(y_i-\bar{Y}_n\right)\left(x_i-\bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i-\bar{X}_n\right)^2} \\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\hat{\alpha}+\hat{\beta}x_i\right)\right]^2 \end{cases}$$

The MLEs for α, β, σ^2 restricted to $\beta \leqslant \beta_0$ are therefore given as:

$$\text{if } \frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2} \leqslant \beta_0, \text{ then: } \begin{cases} \hat{\alpha}_H = \bar{Y}_n - \hat{\beta}_H \bar{X}_n \\ \hat{\beta}_H = \frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2} \\ \hat{\sigma}_H^2 = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\hat{\alpha}_H + \hat{\beta}_H x_i\right)\right]^2 \end{cases}$$

$$\text{otherwise: } \begin{cases} \hat{\alpha}_H = \bar{Y}_n - \hat{\beta}_H \bar{X}_n \\ \hat{\beta}_H = \beta_0 \\ \hat{\sigma}_H^2 = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\hat{\alpha}_H + \hat{\beta}_H x_i\right)\right]^2 \end{cases}$$

The likelihood ratio test is constructed by forming:

$$\Lambda = \frac{L\left(\hat{\alpha}_H, \hat{\beta}_H, \hat{\sigma}_H^2\right)}{L\left(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2\right)}$$

and we reject the null hypothesis H if $\Lambda < c_\alpha \leq 1$[1] for some c_α determined by $P\left(\Lambda < c_\alpha\right) = \alpha$. So, in the case where $\frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2} \leqslant \beta_0$, we will never reject H. Therefore, we construct the numerator of Λ using the restricted MLEs given by:

$$\begin{cases} \hat{\alpha}_H = \bar{Y}_n - \hat{\beta}_H \bar{X}_n \\ \hat{\beta}_H = \beta_0 \\ \hat{\sigma}_H^2 = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\hat{\alpha}_H + \hat{\beta}_H x_i\right)\right]^2 \end{cases}$$

Plugging these values in gives:

$$\Lambda = \frac{L\left(\hat{\alpha}_H, \hat{\beta}_H, \hat{\sigma}_H^2\right)}{L\left(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2\right)} = \frac{\left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha + \beta x_i)\right]^2\right\}\Big|_{(\alpha,\beta,\sigma^2) = \left(\hat{\alpha}_H, \hat{\beta}_H, \hat{\sigma}_H^2\right)}}{\left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[y_i - (\alpha + \beta x_i)\right]^2\right\}\Big|_{(\alpha,\beta,\sigma^2) = \left(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2\right)}}$$

$$= \frac{\left(2\pi\frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\left[\bar{Y}_n - \beta_0 \bar{X}_n\right] + \beta_0 x_i\right)\right]^2\right)^{-\frac{n}{2}}}{\left(2\pi\frac{1}{n}\sum_{i=1}^{n}\left[y_i - \left(\left[\bar{Y}_n - \frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2}\bar{X}\right] + \frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2} x_i\right)\right]^2\right)^{-\frac{n}{2}}}$$

$$= \left(\frac{\sum_{i=1}^{n}\left[y_i - \left(\left[\bar{Y}_n - \beta_0 \bar{X}_n\right] + \beta_0 x_i\right)\right]^2}{\sum_{i=1}^{n}\left[y_i - \left(\left[\bar{Y}_n - b_1 \bar{X}\right] + b_1 x_i\right)\right]^2}\right)^{\frac{n}{2}}$$

where $b_1 = \frac{\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)\left(x_i - \bar{X}_n\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2}$. Manipulation of the expression and corresponding decision rule gives:

$$\left(\frac{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - \beta_0\left(x_i - \bar{X}_n\right)\right]^2}{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - b_1\left(x_i - \bar{X}_n\right)\right]^2}\right)^{\frac{n}{2}} < c_\alpha \Rightarrow \frac{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - \beta_0\left(x_i - \bar{X}_n\right)\right]^2}{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - b_1\left(x_i - \bar{X}_n\right)\right]^2} < c_\alpha^{'}$$

Since $\left(y_i - \bar{Y}_n\right) - \beta_0\left(x_i - \bar{X}_n\right)$ is normally distributed for each i, and it is given that $\frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - b_1\left(x_i - \bar{X}_n\right)\right]^2 \sim \chi^2_{(n-2)}$, the test statistic (with appropriate constant scalings) has the form $\frac{Z^2}{V/D}$, with $Z \sim N(0,1)$ and $V \sim \chi^2_{(d=n-2)}$. This can also be written as:

$$\frac{\chi^2_{(1)}/1}{\chi^2_{(n-2)}/(n-2)} \sim F(1, n-2)$$

in which case our test statistic has an F distribution. Since the numerator and denominator of our test statistic are both positive, we may also compute the square root and obtain a statistic of the form:

$$\frac{Z}{\sqrt{\chi^2_{(n-2)}/(n-2)}} \sim \text{StudentT}(n-2)$$

in which case our test statistic has a Student's T distribution. For this choice, we would reject H in favor of K if $\phi\left(\underset{\sim}{X}, \underset{\sim}{Y}\right) = 1$, where:

$$\phi\left(\underset{\sim}{X}, \underset{\sim}{Y}\right) = \begin{cases} 1 & t < c \\ 0 & o.w. \end{cases}$$

$$t = \frac{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - \beta_0\left(x_i - \bar{X}_n\right)\right]}{\sum_{i=1}^{n}\left[\left(y_i - \bar{Y}_n\right) - b_1\left(x_i - \bar{X}_n\right)\right]}$$

and c is the constant such that $1 - \int_{-\infty}^{c} f_T dt = \alpha$ for some level of significance α with f_T being the density of the Student's T distribution with $n-2$ degrees of freedom. For this problem (as well as the problems which follow) the density function of the test statistic was estimated using a 100,000 iteration (with 15,000 samples each) Monte-Carlo simulation, sorted into 50 categorical bins on the number line.

[1]Here α represents the selected level of significance of the test, not the additive constant regression parameter

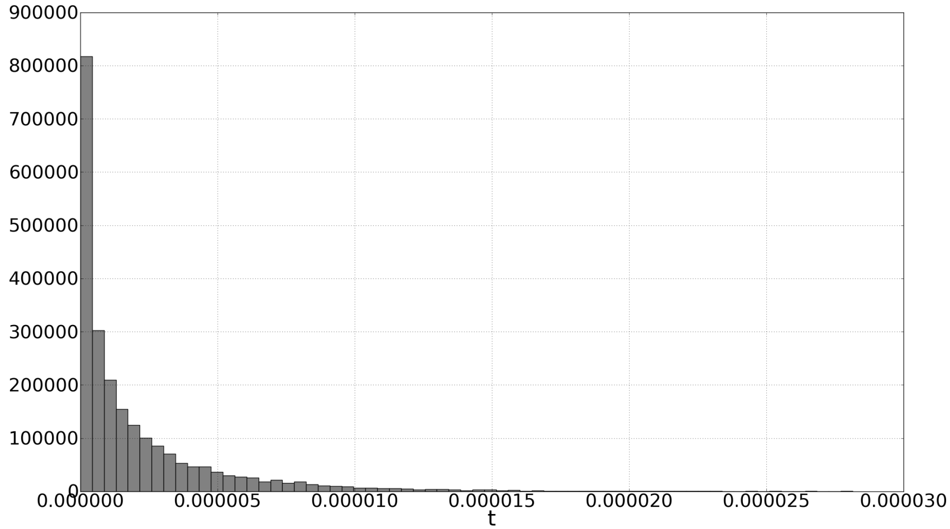

Fig. 1: A plot of the simulated density of the test statistic for Problem 57, without computation of a square-root. This appears to have the shape of an $F-$distribution.

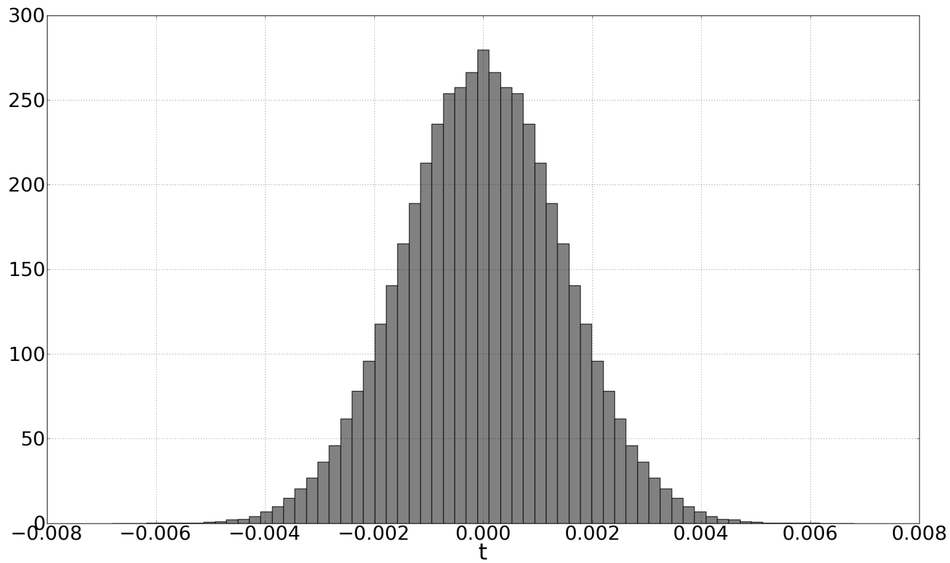

Fig. 2: A plot of the simulated density of the test statistic for Problem 57, with the computation of the square root. For this plot, the histogram was reflected about 0 to induce the proper symmetry. This appears to have the shape of a student's $T-$distribution, as expected.

Problem 3

It is given that $X_1, \dots, X_m \sim N\left(\mu_1, \sigma_1^2\right)$ and $Y_1, \dots, Y_n \sim N\left(\mu_2, \sigma_2^2\right)$, with all parameters unknown. The objective is to formulate a generalized likelihood ratio test for equality of variance:

$$H : \Delta = \Delta_0 \quad vs. \quad K : \Delta \neq \Delta_0$$

with $\Delta = \frac{\sigma_2^2}{\sigma_1^2}$. The likelihood function is given by:

$$L = \left[\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{1}{2\sigma_1^2}\left(x_i - \mu_1\right)^2\right\}\right]\left[\prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{-\frac{1}{2\sigma_2^2}\left(y_j - \mu_2\right)^2\right\}\right]$$

$$= \left(2\pi\sigma_1^2\right)^{-\frac{m}{2}} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{m}\left(x_1 - \mu_1\right)^2\right\}\left(2\pi\sigma_2^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n}\left(y_j - \mu_2\right)^2\right\}$$

$$= \left(2\pi\right)^{-\frac{m+n}{2}}\left(\sigma_1^2\right)^{-\frac{m}{2}}\left(\sigma_2^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{m}\left(x_1 - \mu_1\right)^2 - \frac{1}{2\sigma_2^2}\sum_{j=1}^{n}\left(y_j - \mu_2\right)^2\right\}$$

The log-likelihood function is therefore computed to be:

$$LL = -\frac{m+n}{2}\ln\left(2\pi\right) - \frac{m}{2}\ln\left(\sigma_1^2\right) - \frac{n}{2}\ln\left(\sigma_2^2\right) - \frac{1}{2\sigma_1^2}\sum_{i=1}^{m}\left(x_i - \mu_1\right)^2 - \frac{1}{2\sigma_2^2}\sum_{j=1}^{n}\left(y_j - \mu_2\right)^2$$

Solving for the MLEs $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2$ gives:

$$\frac{\partial LL}{\partial \mu_1} = 0 \Rightarrow \frac{1}{\sigma_1^2}\sum_{i=1}^{m}\left(x_i - \mu_1\right) = 0 \Rightarrow \hat{\mu}_1 = \bar{X}_m$$

$$\frac{\partial LL}{\partial \mu_2} = 0 \Rightarrow \frac{1}{\sigma_2^2}\sum_{i=1}^{n}\left(y_i - \mu_2\right) = 0 \Rightarrow \hat{\mu}_2 = \bar{Y}_n$$

$$\frac{\partial LL}{\partial \sigma_1^2} = 0 \Rightarrow -\frac{m}{2\sigma_1^2} + \frac{1}{2\left(\sigma_1^2\right)^2}\sum_{i=1}^{m}\left(x_i - \mu_1\right)^2 = 0 \Rightarrow \hat{\sigma}_1^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2$$

$$\frac{\partial LL}{\partial \sigma_2^2} = 0 \Rightarrow -\frac{n}{2\sigma_2^2} + \frac{1}{2\left(\sigma_2^2\right)^2}\sum_{i=1}^{n}\left(y_i - \mu_2\right)^2 = 0 \Rightarrow \hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2$$

Under the constraint $\Delta = \frac{\sigma_2^2}{\sigma_1^2} = \Delta_0 \Rightarrow \sigma_2^2 = \Delta_0\sigma_1^2$, the restricted MLEs $\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1^2$ are given as:

$$\frac{\partial LL}{\partial \mu_1} = 0 \Rightarrow \frac{1}{\sigma_1^2}\sum_{i=1}^{m}\left(x_i - \mu_1\right) = 0 \Rightarrow \tilde{\mu}_1 = \hat{\mu}_1 = \bar{X}_m$$

$$\frac{\partial LL}{\partial \mu_2} = 0 \Rightarrow \frac{1}{\Delta_0\sigma_2^2}\sum_{i=1}^{n}\left(y_i - \mu_2\right) = 0 \Rightarrow \tilde{\mu}_2 = \hat{\mu}_2 = \bar{Y}_n$$

$$\frac{\partial LL}{\partial \sigma_1^2} = 0 \Rightarrow \tilde{\sigma}_1^2 = \frac{1}{(m+n)}\left[\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \frac{1}{\Delta_0}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right]$$

The likelihood ratio can then be formed as:

$$\Lambda = \frac{L\left(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2\right)}{L\left(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1^2, \Delta_0\tilde{\sigma}_1^2\right)} = \frac{\left(2\pi\right)^{-\frac{m+n}{2}}\left(\hat{\sigma}_1^2\right)^{-\frac{m}{2}}\left(\hat{\sigma}_2^2\right)^{-\frac{n}{2}}\exp\left\{-\frac{m}{2} - \frac{n}{2}\right\}}{\left(2\pi\right)^{-\frac{m+n}{2}}\left(\tilde{\sigma}_1^2\right)^{-\frac{m+n}{2}}\left(\Delta_0\right)^{-\frac{n}{2}}\exp\left\{-\frac{m}{2} - \frac{n}{2}\right\}}$$

Problem 3 continued on next page. . .

$$= \frac{\left(\tilde{\sigma}_1^2\right)^{\frac{m+n}{2}} \left(\Delta_0\right)^{\frac{n}{2}}}{\left(\hat{\sigma}_1^2\right)^{\frac{m}{2}} \left(\hat{\sigma}_2^2\right)^{\frac{n}{2}}} = \frac{\left(\frac{1}{(m+n)}\left[\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \frac{1}{\Delta_0}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right]\right)^{\frac{m+n}{2}} \left(\Delta_0\right)^{\frac{n}{2}}}{\left(\frac{1}{m}\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{m}{2}} \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{n}{2}}}$$

$$= \frac{\left(\Delta_0 \sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{m+n}{2}} m^{\frac{1}{2}m} n^{\frac{1}{2}n}}{\left(\Delta_0\right)^{\frac{1}{2}m}(m+n)^{\frac{m+n}{2}}\left(\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{1}{2}m}\left(\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{1}{2}n}}$$

The decision rule is to reject H if $\Lambda < c \leq 1$, where c is uniquely determined by $P(\Lambda < c) = \alpha$ for a predetermined significance level α. Simplification by division and multiplication of constant factors gives the rule for rejection:

$$\frac{\left(\Delta_0 \sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{m+n}{2}} m^{\frac{1}{2}m} n^{\frac{1}{2}n}}{\left(\Delta_0\right)^{\frac{1}{2}m}(m+n)^{\frac{m+n}{2}}\left(\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{1}{2}m}\left(\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{1}{2}n}} < c$$

$$\Rightarrow \frac{\left(\Delta_0 \sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{m+n}{2}}}{\left(\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{1}{2}m}\left(\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{1}{2}n}} < c'$$

$$\Rightarrow \frac{\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \frac{1}{\Delta_0}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2}{\left(\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{m}{m+n}}\left(\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{n}{m+n}}} < c''$$

So, our test function is given by:

$$\phi(Z_1, \ldots, Z_n) = \begin{cases} 1 & \frac{\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2 + \frac{1}{\Delta_0}\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2}{\left(\sum_{i=1}^{m}\left(x_i - \bar{X}_m\right)^2\right)^{\frac{m}{m+n}}\left(\sum_{i=1}^{n}\left(y_i - \bar{Y}_n\right)^2\right)^{\frac{n}{m+n}}} < c, \text{ for } c : \int_{-\infty}^{c} f_t dt = \alpha \\ 0 & o.w. \end{cases}$$

The density of the test statistic is estimated in Fig. 3.

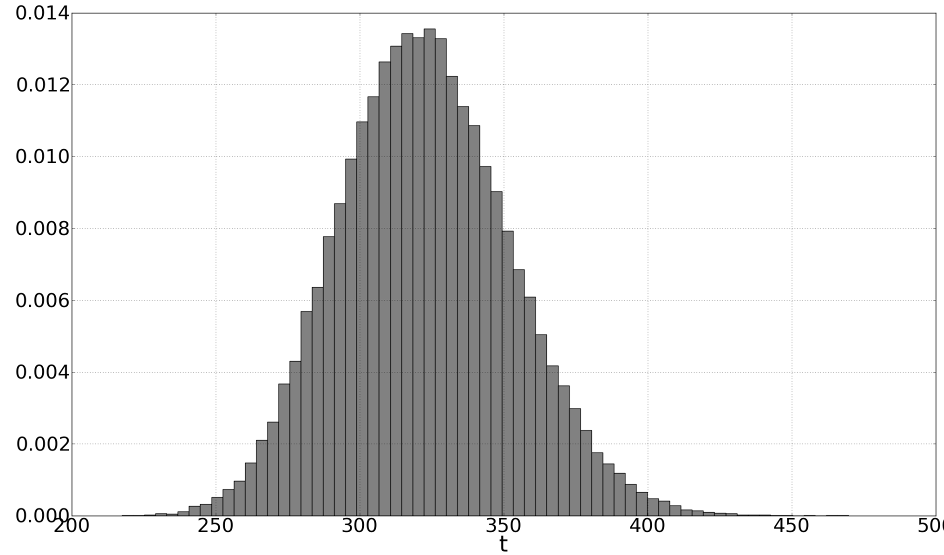

Fig. 3: A plot of the simulated density of the test statistic for Problem 58. This density appears to be non-negative and slightly right-skew; similar to a Gamma distribution.

Problem 4

It is given that $Z_1, \ldots, Z_n \sim N\left(\mu_i, \sigma^2\right)$, with $\mu_{i>s} = 0$ for some fixed known number $0 < s < n$. It is required to find a generalized likelihood ratio test for:

$$H : \mu_1 \leqslant \mu_1^0 \quad \text{vs.} \quad K : \mu_1 > \mu_1^0$$

The likelihood function is computed as:

$$L = \left[\prod_{i=1}^{s} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(z_i - \mu_i\right)^2\right\}\right]\left[\prod_{j=s+1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} z_j^2\right\}\right]$$

$$= \left(2\pi\sigma^2\right)^{-\frac{s}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{s}\left(z_i - \mu_i\right)^2\right\}\left(2\pi\sigma^2\right)^{-\frac{n-s}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=s+1}^{n} z_j^2\right\}$$

$$= \left(2\pi\sigma^2\right)^{-\frac{s}{2}-\frac{n-s}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{s}\left(z_i - \mu_i\right)^2 - \frac{1}{2\sigma^2}\sum_{j=s+1}^{n} z_j^2\right\}$$

$$= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{s}\left(z_i - \mu_i\right)^2 + \sum_{j=s+1}^{n} z_j^2\right]\right\}$$

The log-likelihood function is therefore:

$$LL = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\sigma^2\right) - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{s}\left(z_i - \mu_i\right)^2 + \sum_{j=s+1}^{n} z_j^2\right]$$

The MLEs $\hat{\mu}_1, \ldots, \hat{\mu}_s, \hat{\sigma}^2$ are therefore determined by:

$$\frac{\partial LL}{\partial \mu_1} = 0 \Rightarrow -2z_1 - 2\mu_1 = 0 \Rightarrow \hat{\mu}_1 = z_1$$

$$\vdots$$

$$\frac{\partial LL}{\partial \mu_s} = 0 \Rightarrow -2z_s - 2\mu_s = 0 \Rightarrow \hat{\mu}_s = z_s$$

$$\frac{\partial LL}{\partial \sigma^2} = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{s} (z_i - \hat{\mu}_i)^2 + \sum_{j=s+1}^{n} z_j^2}{n} = \frac{1}{n}\sum_{j=s+1}^{n} z_j^2$$

The MLEs $\tilde{\mu}_1, \ldots, \tilde{\mu}_s, \tilde{\sigma}^2$ subject to the constraint $\mu_1 \leqslant \mu_1^0$ are determined by:

$$\frac{\partial LL}{\partial \mu_1} = 0 \Rightarrow -2z_1 - 2\mu_1 = 0, \mu_1 \leqslant \mu_1^0 \Rightarrow \tilde{\mu}_1 = \mu_1^0$$

$$\frac{\partial LL}{\partial \mu_2} = 0 \Rightarrow -2z_2 - 2\mu_2 = 0 \Rightarrow \tilde{\mu}_2 = z_2$$

$$\vdots$$

$$\frac{\partial LL}{\partial \mu_s} = 0 \Rightarrow -2z_s - 2\mu_s = 0 \Rightarrow \tilde{\mu}_s = z_s$$

$$\frac{\partial LL}{\partial \sigma^2} = 0 \Rightarrow \tilde{\sigma}^2 = \frac{\sum_{i=1}^{s} (z_i - \hat{\mu}_i)^2 + \sum_{j=s+1}^{n} z_j^2}{n} = \frac{1}{n}\left[\left(z_1 - \mu_1^0\right)^2 + \sum_{j=s+1}^{n} z_j^2\right]$$

Since we are interested in establishing the criteria for rejection of H, we disregard the situations where $\mu_1^0 > z_1 \Rightarrow \tilde{\mu}_1 = z_1$; since this will lead to $\Lambda = 1$ and we would fail to reject H for any such case. Forming the likelihood ratio gives us:

$$\Lambda = \frac{L\left(\tilde{\mu}_1, \ldots, \tilde{\mu}_s, \tilde{\sigma}^2\right)}{L\left(\hat{\mu}_1, \ldots, \hat{\mu}_s, \hat{\sigma}^2\right)} = \frac{\left(2\pi \frac{1}{n}\left[\left(z_1 - \mu_1^0\right)^2 + \sum_{j=s+1}^{n} z_j^2\right]\right)^{-\frac{n}{2}} \exp\left\{-\frac{n}{2}\right\}}{\left(2\pi \frac{1}{n}\sum_{j=s+1}^{n} z_j^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{n}{2}\right\}}$$

$$= \frac{\left(\left(z_1 - \mu_1^0\right)^2 + \sum_{j=s+1}^{n} z_j^2\right)^{-\frac{n}{2}}}{\left(\sum_{j=s+1}^{n} z_j^2\right)^{-\frac{n}{2}}} = \left(\frac{\left(z_1 - \mu_1^0\right)^2}{\sum_{j=s+1}^{n} z_j^2} + 1\right)^{-\frac{n}{2}}$$

The decision rule is to reject H if $\Lambda < c \leq 1$, where c is uniquely determined by $P\left(\Lambda < c\right) = \alpha$ for a predetermined significance level α. Simplification by division and multiplication of constant factors gives the rule for rejection:

$$\left(\frac{\left(z_1 - \mu_1^0\right)^2}{\sum_{j=s+1}^{n} z_j^2} + 1\right)^{-\frac{n}{2}} < c \Rightarrow \frac{\left(z_1 - \mu_1^0\right)^2}{\sum_{j=s+1}^{n} z_j^2} + 1 < c' \Rightarrow \frac{\left(z_1 - \mu_1^0\right)^2}{\sum_{j=s+1}^{n} z_j^2} < c''$$

Since $z_1 - \mu_1^0$ and each z_j is normally distributed, our test statistic is recognized as having a T distribution (through proper normalization):

$$\frac{\left(z_1 - \mu_1^0\right)^2}{\frac{1}{n-s}\sum_{j=s+1}^{n} z_j^2} \to \frac{z_1 - \mu_1^0}{\sqrt{\frac{1}{n-s-1}\sum_{j=s+1}^{n} z_j^2}} \to \frac{Z}{\sqrt{V/d}} \sim T\left(n-s-1\right)$$

Therefore, our test is to reject H in favor of K if $\phi\left(Z_1, \ldots, Z_n\right) = 1$, for ϕ defined as:

$$\phi\left(Z_1, \ldots, Z_n\right) = \begin{cases} 1 & \frac{z_1 - \mu_1^0}{\sqrt{\frac{1}{n-s-1}\sum_{j=s+1}^{n} z_j^2}} < c, \text{ for } c \text{ such that } 1 - \int_{\infty}^{c} f_T dt = \alpha \\ 0 & o.w. \end{cases}$$

where f_T is the density of the Student's T distribution with $n-s-1$ degrees of freedom, and α is the chosen level of significance for the test. The density of the test statistic is estimated in Fig. 4.

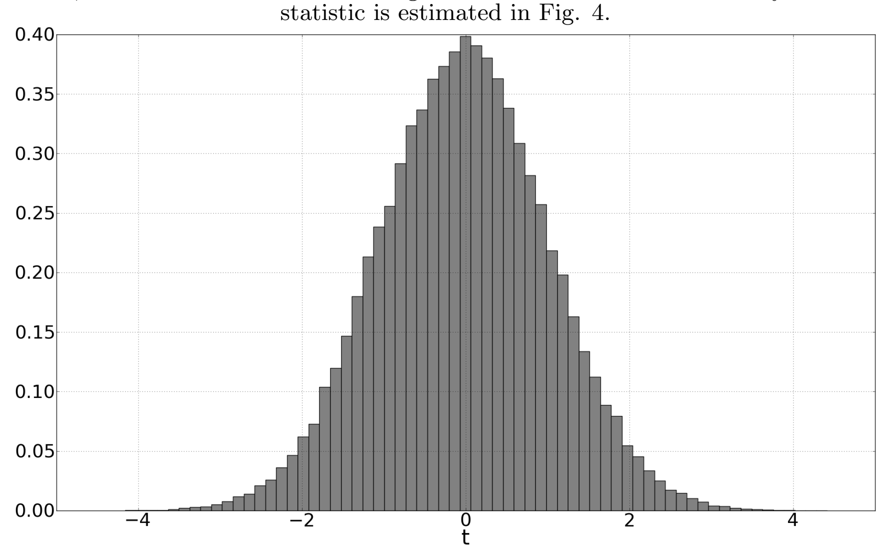

Fig. 4: A plot of the simulated density of the test statistic for Problem 59. As expected, this appears to have the shape of a T−distribution.

Problem 5

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} Exp\left(\lambda\right)$. It is required to find a UMP test for:

$$H : \lambda \leqslant 1 \quad \text{vs.} \quad K : \lambda > 1$$

First, consider the formulation of a generalized likelihood ratio test for the hypotheses:

$$H' : \lambda = \lambda_0 \quad \text{vs.} \quad K' : \lambda = \lambda_1$$

Λ is computed to be:

$$\Lambda = \frac{L(\lambda_0)}{L(\lambda_1)} = \frac{\prod_{i=1}^{n} \lambda_0 \exp\{-\lambda_0 x_i\}}{\prod_{i=1}^{n} \lambda_1 \exp\{-\lambda_1 x_i\}} = \frac{(\lambda_0)^n \exp\left\{-\lambda_0 \sum_{i=1}^{n} x_i\right\}}{(\lambda_1)^n \exp\left\{-\lambda_1 \sum_{i=1}^{n} x_i\right\}}$$

$$= (\lambda_0\lambda_1)^n \exp\left\{-\lambda_0 \sum_{i=1}^{n} x_i + \lambda_1 \sum_{i=1}^{n} x_i\right\} = (\lambda_0\lambda_1)^n \exp\left\{(\lambda_1 - \lambda_0)\sum_{i=1}^{n} x_i\right\}$$

The decision rule is to reject H if $\Lambda < c \leq 1$, where c is uniquely determined by $P(\Lambda < c) = \alpha = 0.05$. Simplification by division and multiplication of constant factors gives the rule for rejection:

$$(\lambda_0\lambda_1)^n \exp\left\{(\lambda_1 - \lambda_0)\sum_{i=1}^{n} x_i\right\} < c \Rightarrow \exp\left\{(\lambda_1 - \lambda_0)\sum_{i=1}^{n} x_i\right\} < c'$$

If we assume that $\lambda_1 > \lambda_0$:

$$\Rightarrow (\lambda_1 - \lambda_0)\sum_{i=1}^{n} x_i < c'' \Rightarrow \sum_{i=1}^{n} x_i < c'''$$

The test statistic $\sum_{i=1}^{n} x_i$ is recognized as having a Gamma(n, λ_0) distribution. Therefore, our test function will have the form:

$$\phi(Z_1, \ldots, Z_n) = \begin{cases} 1 & \sum_{i=1}^{n} x_i < c, \text{ for } c \text{ such that } \int_{-\infty}^{c} f_{\Gamma(n,\lambda_0)} dt = 0.05 \\ 0 & o.w. \end{cases}$$

where $f_{\Gamma(n,\lambda_0)}$ is the density function of the Gamma distribution with parameters n and λ_0. The inequality will change direction if $\lambda_1 < \lambda_0$. Other than this, it is noted that the test functions do not depend on the exact value of λ_1. Therefore, this test is most powerful by the Neyman-Pearson Lemma for tests involving alternative hypothesis of the form:

$$H^* : \lambda = 1 \quad \text{vs.} \quad K^* : \lambda > 1$$

Now we want to show that the power function $\pi(\phi, \lambda) = P\left(\sum_{i=1}^{n} x_i < c\right)$ is monotonic in λ. Since:

$$z \sim \exp(\lambda) \Rightarrow \lambda u \sim \exp(\lambda) \Rightarrow u \sim \frac{1}{\lambda}\exp(1)$$

and a sum of exponential distributions has a Gamma distribution, we have:

$$\max_{x \leqslant 1} P\left(\sum_{i=1}^{n} \frac{1}{\lambda} x_i < c\right) \Rightarrow \max_{x \leqslant 1} P \underbrace{\left(\sum_{i=1}^{n} x_i < \lambda c\right)}_{P\nearrow \text{ as } \lambda\nearrow;\text{ so monotonic}}$$

Due to monotonicity, we may conclude that the UMP test for:

$$H : \lambda \leqslant 1 \quad \text{vs.} \quad K : \lambda > 1$$

is specified by the test function (decision rule):

$$\phi(Z_1, \ldots, Z_n) = \begin{cases} 1 & \sum_{i=1}^{n} x_i < c, \text{ for } c \text{ such that } \int_{-\infty}^{c} f_{\Gamma(1,n)} dt = 0.05 \\ 0 & o.w. \end{cases}$$

A plot of the estimated distribution of the test statistic is given in Fig. 5. The estimated value of c is computed from the exact PDF of a Gamma$(1, n)$ distribution; shown in Fig.6.

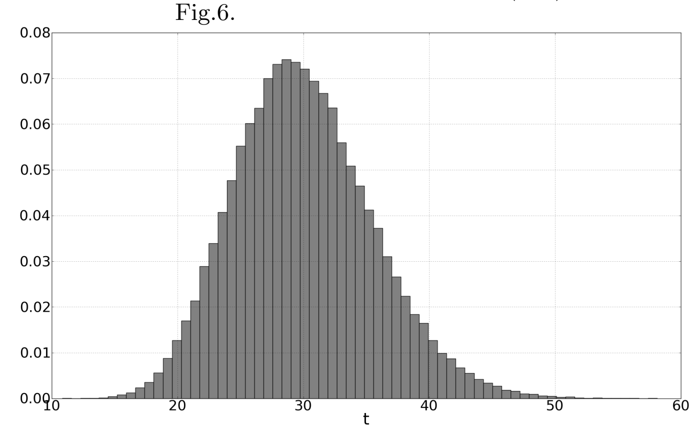

Fig. 5: A plot of the estimated density of the test statistic for Problem 60; assuming $n = 30$. This appears similar to the Gamma distribution with parameters 1 and $n = 30$, as expected.

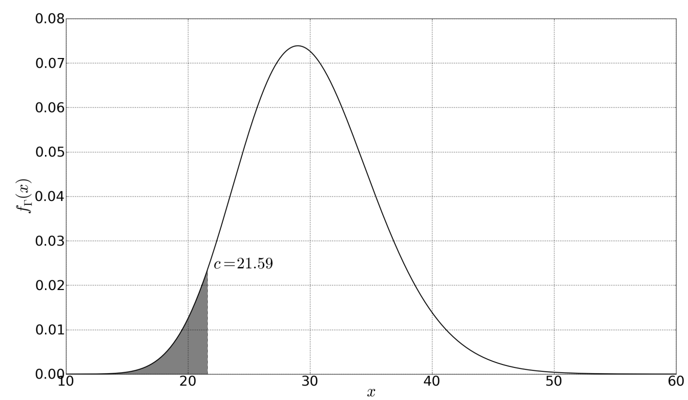

Fig. 6: A plot of the exact density of the test statistic for Problem 60; assuming $n = 30$. This is given by the PDF of $\Gamma(1, n)$. The corresponding value of $c \approx 21.59$ for $\alpha = 0.05$ is also shown.