STAT 305 Assignment 4

Name: Kevin Nguyen
Student Number: 97703045

Question 1

1. We have that $Z = \sum_{i=0}^{n} Y_i \sim N(n\mu, n\sigma^2)$ due to normality of sum of normal distribution. The likelihood ratio is

$$
\begin{aligned}
LR &= \frac{f_Z(z \mid \sigma^2 = \sigma_a^2)}{f_Z(z \mid \sigma^2 = \sigma_0^2)} \\
&= \frac{\frac{1}{(2\pi)^{n/2}} \left(\frac{1}{\sigma_a^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma_a^2} \sum_{i=1}^{n} (y_i - \mu)^2\right)}{\frac{1}{(2\pi)^{n/2}} \left(\frac{1}{\sigma_0^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (y_i - \mu)^2\right)} \\
&= \left(\frac{\sigma_0^2}{\sigma_a^2}\right)^{n/2} \exp\left[\left(-\frac{1}{2\sigma_a^2} + \frac{1}{2\sigma_0^2}\right) \sum_{i=1}^{n} (y_i - \mu)^2\right]
\end{aligned}
$$

2. This is a decreasing function because the fraction before the exponential is a constant. Furthermore, $\left(-\frac{1}{2\sigma_a^2} + \frac{1}{2\sigma_0^2}\right)$ will be negative since $\sigma_a^2 < \sigma_0^2$ and e^{-x} is a decreasing function. Therefore, the likelihood ratio is a decreasing function

3. Under H_0, we have $\sigma^2 = \sigma_0^2$. We also have that

$$
\begin{aligned}
\frac{\sum_{i=1}^{n} (Y_i - \mu)^2}{\sigma_0^2} &= \sum_{i=1}^{n} \frac{(Y_i - \mu)^2}{\sigma_0^2} \\
&= \sum_{i=1}^{n} \left(\frac{Y_i - \mu}{\sigma_0}\right)^2 \\
&= \sum_{i=1}^{n} Z^2 && \text{where } Z \sim N(0,1) \text{ because of standardization} \\
&= \sum_{i=1}^{n} S && \text{where } S \sim \chi_1^2 \text{ via distribution of square of standard normal} \\
&= F && \text{where } F \sim \chi_n^2 \text{ via sum of Chi-squared random variables}
\end{aligned}
$$

Therefore, under the null, $\frac{\sum_{i=1}^{n} (Y_i - \mu)^2}{\sigma_0^2}$ follows the chi squared distribution with n-degrees of freedom. This is an exact result.

4. Let $T = \left[\sum_{i=1}^{n}(Y_i - \mu)^2\right]/\sigma^2$. We are looking for c such that

$$P(T < c \mid \sigma^2 = \sigma_0^2) = \alpha \quad \text{where } c = qchisq(\alpha, n)$$

So c is the quantile of a chi-square distribution with n degree of freedom. Therefore, we reject when the test statistic is less than c.

5. For this question, we have

$$\begin{aligned}
T &= \frac{\sum_{i=1}^{50}(Y_i - \mu)^2}{\sigma_0^2} \\
&= \frac{\sum_{i=1}^{50} Y_i^2 - 2Y_i\mu + \mu^2}{25} \\
&= \frac{\sum_{i=1}^{50} Y_i^2 - \sum_{i=1}^{50} 2Y_i\mu + \sum_{i=1}^{50} \mu^2}{25} \\
&= \frac{\sum_{i=1}^{50} Y_i^2 - 2\mu\sum_{i=1}^{50} Y_i + \sum_{i=1}^{n} \mu^2}{25} \\
&= \frac{200 - 2(1)(60) + 50}{25} \\
&= 5.2
\end{aligned}$$

$$\begin{aligned}
\text{when } \alpha = 0.05, \quad c_{0.05} &= qchisq(\alpha, n) \\
&= qchisq(0.05, 50) \\
&= 34.76425
\end{aligned}$$

$$\begin{aligned}
\text{when } \alpha = 0.01, \quad c_{0.01} &= sqchisq(0.01, 1) \\
&= 29.70668
\end{aligned}$$

Therefore, at both significance level, we do not reject H_0

Question 2

1. Similarly to above, we have

$$\begin{aligned}
LR &= \frac{f_{Y_1,\ldots,Y_n(y_1,\ldots,y_n,|\mu_a,\sigma^2)}}{f_{Y_1,\ldots,Y_n(y_1,\ldots,y_n,|\mu_0,\sigma^2)}} \\
&= \frac{\frac{1}{(2\pi)^{n/2}}\left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(y_i-\mu_a)^2\right)}{\frac{1}{(2\pi)^{n/2}}\left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(y_i-\mu_0)^2\right)} \\
&= \frac{\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(y_i-\mu_a)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(y_i-\mu_0)^2\right)} \\
&= \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n(y_i-\mu_a)^2-\sum_{i=1}^n(y_i-\mu_0)^2\right)\right] \\
\text{note } &: \sum_{i=1}^n(y_i-\mu_{a/0})^2 = \sum_{i=1}^n(y_i-\bar{y}+\bar{y}-\mu_{a/0}) = \sum_{i=1}^n(y_i-\bar{y})^2+n(\bar{y}-\mu_{a/0})^2 \\
&= \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n(y_i-\bar{y})^2+n(\bar{y}-\mu_a)^2-\sum_{i=1}^n(y_i-\bar{y})^2-n(\bar{y}-\mu_0)^2\right)\right] \\
&= \exp\left[-\frac{1}{2\sigma^2}\left(n(\bar{y}-\mu_a)^2-n(\bar{y}-\mu_0)^2\right)\right] \\
&= \exp\left[-\frac{n}{2\sigma^2}\left[(\mu_a^2-\mu_0^2)-2\bar{y}(\mu_a-\mu_0)\right]\right] \\
&= \exp\left[-\frac{n}{2\sigma^2}(\mu_a^2-\mu_0^2)\right]\exp\left[\frac{n\bar{y}}{\sigma^2}(\mu_a-\mu_0)\right]
\end{aligned}$$

   And since $\mu_a > \mu_0$ (since $\mu_0 = 0$ and $\mu_a > 0$), the likelihood ratio is a monotonically increasing function of $\bar{y}$ thus we have

$$\frac{f_{Y_1,\ldots,Y_n(y_1,\ldots,y_n,|\mu_a,\sigma^2)}}{f_{Y_1,\ldots,Y_n(y_1,\ldots,y_n,|\mu_0,\sigma^2)}} > c$$

   Which is equivalent to $\bar{y} > c'$ where c and c' are some constants

2. As proven above, as long as we have $\mu_a > \mu_0 = 0$, we will the fact that the LR is monotonically increasing function of $\bar{y}$ and thus even though the the decision (rejection) region might change, the decision rule will remain that same with respect to the sample mean $\bar{y}$ and some constant c'

3. Let $Y_1, \ldots, Y_n$ be a sample of n independent observations from a $N(\mu, \sigma^2)$ distribution as stated in the question. We then have

$$\frac{\bar{Y} - \mu}{\sqrt{S^2/n}} \sim t_{n-1}$$

where both S is the sample variance, note here that both Y and S are random variables.

Just as above where we had the test statistics $T_{\text{known variance}} = \bar{Y} \sim N(n\mu, n\sigma)$. Since we do not know the variance, we are using the t-test.

When the variance was known, our decision rule was to reject if $T > c'_{\text{known variance}}$ and $c'_{\text{known variance}} = \text{qnorm}(1 - \alpha, n\mu, \sqrt{n\sigma})$. We can apply the same logic to the unknown variance case and say

$$\text{Reject when } T > c' \qquad \text{where } c' = \text{qt}(1 - \alpha, n - 1)$$

in other words, we reject when $T > t_{n-1,1-\alpha}$