1. (a) Firstly, the (log-)likelihood of observing $(X_1, \cdots, X_n)$ in terms of (a, θ) is

$$
\begin{aligned}
L(a, \theta|\mathbf{x}) &= \left(\frac{1}{\sqrt{2\pi a\theta}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2a\theta}\right\} \\
l(a, \theta|\mathbf{x}) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(a\theta) - \frac{\sum_{i=1}^n (x_i - \theta)^2}{2a\theta},
\end{aligned}
$$

whose first-order conditions are

$$
\begin{cases}
\dfrac{\partial l(a, \theta)}{\partial a} = -\dfrac{n}{2a} + \dfrac{\sum_{i=1}^n (x_i - \theta)^2}{2a^2\theta} = 0 \\[2ex]
\dfrac{\partial l(a, \theta)}{\partial \theta} = -\dfrac{n}{2\theta} + \dfrac{\sum_{i=1}^n x_i^2}{2a\theta^2} - \dfrac{n}{2a} = 0
\end{cases}.
$$

Solving yields

$$
\begin{cases}
\theta = \dfrac{\sum_{i=1}^n x_i}{n} = \bar{x} \\[2ex]
a = \dfrac{\sum_{i=1}^n (x_i - \bar{x})^2}{n\bar{x}} = \dfrac{s_n^2}{\bar{x}}
\end{cases},
$$

where $s_n^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$. By further observing the concavity of $l(a, \theta)$, we have $(\hat{a}, \hat{\theta}) = (\bar{x}, \frac{s_n^2}{\bar{x}})$ being the MLE of unconstrained (a, θ), i.e.

$$
\sup_{(a,\theta)\in\Theta} L(a, \theta|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\hat{a}\hat{\theta}}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i - \hat{\theta})^2}{2\hat{a}\hat{\theta}}\right\} = \left(\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi s_n^2}}\right)^n.
$$

Under H_0, the MLE for θ can also be obtained through the first-order condition (and concavity)

$$
-\frac{n}{2\theta} + \frac{\sum_{i=1}^n x_i^2}{2\theta^2} - \frac{n}{2} = 0 \implies \hat{\theta}_0 = \frac{-1 + \sqrt{1 + 4n^{-1}\sum_{i=1}^n x_i^2}}{2},
$$

indicating that

$$
\sup_{(a,\theta)\in\Theta_0} L(a, \theta|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\hat{\theta}_0}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i - \hat{\theta}_0)^2}{2\hat{\theta}_0}\right\}.
$$

Hence, the required LRT is to reject H_0 if the likelihood ratio

$$
\begin{aligned}
\lambda(\mathbf{x}) &= \frac{\sup_{(a,\theta)\in\Theta_0} L(a, \theta|\mathbf{x})}{\sup_{(a,\theta)\in\Theta} L(a, \theta|\mathbf{x})} \\
&= \left(\frac{s_n^2}{\hat{\theta}_0}\right)^{\frac{n}{2}} \exp\left\{\frac{n}{2} - \frac{\sum_{i=1}^n (x_i - \hat{\theta}_0)^2}{2\hat{\theta}_0}\right\}
\end{aligned}
$$

is reasonably small, i.e. $\lambda(\mathbf{x}) < c$ for some c.

(b) We copy and paste and modify from (a). The (log-)likelihood of observing $(X_1, \cdots, X_n)$ is

$$
\begin{aligned}
L(a, \theta|\mathbf{x}) &= \left(\frac{1}{\sqrt{2\pi a\theta^2}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2a\theta^2}\right\} \\
l(a, \theta|\mathbf{x}) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(a\theta^2) - \frac{\sum_{i=1}^n (x_i - \theta)^2}{2a\theta^2},
\end{aligned}
$$

whose first-order conditions are

$$\begin{cases} \dfrac{\partial l(a,\theta)}{\partial a} = -\dfrac{n}{2a} + \dfrac{\sum_{i=1}^n (x_i-\theta)^2}{2a^2\theta^2} = 0 \\ \\ \dfrac{\partial l(a,\theta)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{\sum_{i=1}^n x_i^2}{a\theta^3} - \dfrac{\sum_{i=1}^n x_i}{a\theta^2} = 0 \end{cases}.$$

Solving yields

$$\begin{cases} \theta = \dfrac{\sum_{i=1}^n x_i}{n} = \bar{x} \\ \\ a = \dfrac{\sum_{i=1}^n (x_i-\bar{x})^2}{n\bar{x}^2} = \dfrac{s_n^2}{\bar{x}^2} \end{cases}.$$

By further observing the concavity of $l(a,\theta)$, we have $(\hat{a},\hat{\theta}) = (\bar{x}, \frac{s_n^2}{\bar{x}^2})$ being the MLE of unconstrained (a,θ), i.e.

$$\sup_{(a,\theta)\in\Theta} L(a,\theta|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\hat{a}\hat{\theta}^2}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i-\hat{\theta})^2}{2\hat{a}\hat{\theta}^2}\right\} = \left(\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi s_n^2}}\right)^n.$$

Under H_0, the MLE for θ can also be obtained through the first-order condition (and concavity)

$$-\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i^2}{\theta^3} - \frac{\sum_{i=1}^n x_i}{\theta^2} = 0 \implies \hat{\theta}_0 = \frac{-\bar{x} + \sqrt{\bar{x}^2 + 4n^{-1}\sum_{i=1}^n x_i^2}}{2},$$

indicating that

$$\sup_{(a,\theta)\in\Theta_0} L(a,\theta|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\hat{\theta}_0^2}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i-\hat{\theta}_0)^2}{2\hat{\theta}_0^2}\right\}.$$

Hence, the required LRT is to reject H_0 if the likelihood ratio

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{\sup_{(a,\theta)\in\Theta_0} L(a,\theta|\mathbf{x})}{\sup_{(a,\theta)\in\Theta} L(a,\theta|\mathbf{x})} \\ &= \left(\frac{s_n^2}{\hat{\theta}_0^2}\right)^{\frac{n}{2}} \exp\left\{\frac{n}{2} - \frac{\sum_{i=1}^n (x_i-\hat{\theta}_0)^2}{2\hat{\theta}_0^2}\right\} \end{aligned}$$

is reasonably small.

2. Note that the likelihood ratio for simple H_0 against simple H_1 is

$$\lambda(\mathbf{x}) = \frac{L(\sigma_1|\mathbf{x})}{L(\sigma_0|\mathbf{x})} = \frac{\left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)^n \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^n x_i^2\right\}}{\left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2\right\}} = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left\{\frac{1}{2}\sum_{i=1}^n x_i^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\right\}.$$

By Neyman-Pearson Lemma, the most powerful test is

$$\begin{aligned}
\psi^* &= \mathbb{1}\left\{\lambda(\mathbf{X}) = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left\{\frac{1}{2}\sum_{i=1}^n X_i^2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right)\right\} > k\right\} \\
&= \mathbb{1}\left\{\exp\left\{\frac{1}{2}\sum_{i=1}^n X_i^2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right)\right\} > k\left(\frac{\sigma_1}{\sigma_0}\right)^n\right\} \\
&= \mathbb{1}\left\{\sum_{i=1}^n X_i^2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right) > 2\log\left(k\left(\frac{\sigma_1}{\sigma_0}\right)^n\right)\right\} \\
&= \mathbb{1}\left\{\sum_{i=1}^n X_i^2 > 2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right)^{-1}\log\left(k\left(\frac{\sigma_1}{\sigma_0}\right)^n\right)\right\} =: \mathbb{1}\left\{\sum_{i=1}^n X_i^2 > c\right\},
\end{aligned}$$

as desired. Here we used the fact that $\sigma_1 > \sigma_0 > 0$ so that the inequality sign is preserved.

If H_0 is true, then we have $V := \frac{\sum_{i=1}^n X_i^2}{\sigma_0^2} \sim \chi_n^2$ and $\mathbb{P}(\sum_{i=1}^n X_i^2 > c) = \mathbb{P}(V > \frac{c}{\sigma_0^2}) = \alpha$, hence it follows that $\frac{c}{\sigma_0^2} = \chi_n^2(1-\alpha)$, or $c = \sigma_0^2\chi_n^2(1-\alpha)$, where $\chi_n^2(1-\alpha)$ is the $(1-\alpha)$-th quantile of the pdf of χ_n^2.

3. (a) If $H_0 : \theta = 0$ is true, then we should have

$$\begin{aligned}
\alpha &= \mathbb{P}(X_{(n)} \geq 1 \text{ or } X_{(1)} > k) = \mathbb{P}(X_{(1)} > k) \\
&= \mathbb{P}(X_i > k \text{ for } i = 1, \cdots, n) = [\mathbb{P}(X_1 > k)]^n = (1-k)^n.
\end{aligned}$$

Thus, $k = 1 - \alpha^{\frac{1}{n}}$.

(b) By using the fact that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$, we have

$$\begin{aligned}
\beta(\theta) &= \mathbb{P}_\theta(X_{(n)} \geq 1 \text{ or } X_{(1)} > k) \\
&= 1 - \mathbb{P}_\theta(X_{(n)} < 1) + \mathbb{P}_\theta(X_{(1)} > k, X_{(n)} < 1) \\
&= 1 - [\mathbb{P}_\theta(X_1 < 1)]^n + [\mathbb{P}_\theta(k < X_1 < 1)]^n.
\end{aligned}$$

Since $1 > k > 0$ for $\alpha > 0$, we can determine $\beta(\theta)$ by partitioning $\{\theta \in \mathbb{R}\}$ into the following cases:

1. $\theta < \theta + 1 < k < 1$. Then $\beta(\theta) = 1 - 1 + 0 = 0$.
2. $\theta < k \leq \theta + 1 < 1$. Then $\beta(\theta) = 1 - 1 + [\mathbb{P}(X_1 > k)]^n = (1 + \theta - k)^n$.
3. $\theta < k < 1 \leq \theta + 1$. Then $\beta(\theta) = 1 - (1-\theta)^n + (1-k)^n = 1 + \alpha - (1-\theta)^n$.
4. $k \leq \theta < \theta + 1$. Then $\beta(\theta) = 1 - [\mathbb{P}_\theta(X_1 < 1)]^n + [\mathbb{P}_\theta(X_1 < 1)]^n = 1$.

If the parameter space is $\{\theta \geq 0\}$ only, then cases 1 and 2 are out of our interest.

(c) By the lemma on Page 23 of Lecture Note 5, it suffices to show that whenever we fix an arbitrary $\theta_1 > 0$, the test is the most powerful one for testing $H_0 : \theta = 0$ against $H_1 : \theta = \theta_1$. Now, the likelihood ratio here is

$$\lambda(\mathbf{x}) = \frac{L(\theta_1|\mathbf{x})}{L(0|\mathbf{x})} = \frac{\mathbb{1}\{\theta_1 < x_{(1)}, x_{(n)} < \theta_1 + 1\}}{\mathbb{1}\{0 < x_{(1)}, x_{(n)} < 1\}},$$

and we proceed by splitting $\{\theta_1 > 0\}$ into the following three cases. By Neyman-Pearson Lemma, it remains to find a threshold c in each case, such that both (i) $\lambda(\mathbf{x}) > c \implies x_{(n)} \geq 1$ or $x_{(1)} > k$, and (ii) $\lambda(\mathbf{x}) < c \implies x_{(n)} < 1$ and $x_{(1)} \leq k$ holds.

1. $0 < \theta_1 < k < 1$. Then

$$\lambda(\mathbf{x}) = \begin{cases} 0, & \text{if } 0 < x_{(1)} < \theta_1, x_{(n)} < 1 \\ 1, & \text{if } \theta_1 < x_{(1)}, x_{(n)} < 1 \\ \infty, & \text{if } \theta_1 < x_{(1)}, 1 \leq x_{(n)} < \theta_1 + 1 \end{cases}.$$

Pick $c = 1$ so we have

$$\begin{aligned} \lambda(\mathbf{x}) > c &\implies x_{(n)} \geq 1, \\ \lambda(\mathbf{x}) < c &\implies x_{(n)} < 1 \text{ and } x_{(1)} < \theta_1 < k. \end{aligned}$$

2. $0 < k \leq \theta_1 < 1$. Then $\lambda(\mathbf{x})$ is the same as above, but we pick $c = 0$, which yields

$$\lambda(\mathbf{x}) > c \implies x_{(1)} > \theta_1 \geq k,$$

and $\lambda(\mathbf{x}) < c$, which is a false statement, implies anything.

3. $0 < k < 1 \leq \theta_1$. Then

$$\lambda(\mathbf{x}) = \begin{cases} 0, & \text{if } 0 < x_{(1)}, x_{(n)} < 1 \\ \infty, & \text{if } \theta_1 < x_{(1)}, x_{(n)} < \theta_1 + 1 \end{cases}.$$

Pick $c = 0$ and reuse the arguments in case 2.

4. Referring to Definition 8.3.16 of Casella & Burger, we assume that "has MLR" means "has MLR in x", i.e. the (log-)likelihood ratio $\log f_{\theta_2}(x) - \log f_{\theta_1}(x)$ is non-decreasing in x for all $\theta_1 < \theta_2$. This is equivalent to

$$0 \leq \frac{\partial}{\partial x} (\log f_{\theta_2}(x) - \log f_{\theta_1}(x)) = \frac{\partial}{\partial x} \log f_{\theta_2}(x) - \frac{\partial}{\partial x} \log f_{\theta_1}(x)$$

for all x and $\theta_1 < \theta_2$. But immediately from definition of non-decreasing, the inequality indicates that $\frac{\partial}{\partial x} \log f_\theta(x)$ is non-decreasing in θ for all x. This is equivalent to

$$\begin{aligned} 0 &\leq \frac{\partial}{\partial \theta} \frac{\partial}{\partial x} \log f_\theta(x) \\ &= \frac{\partial}{\partial \theta} \frac{\frac{\partial}{\partial x} f_\theta(x)}{f_\theta(x)} = \frac{f_\theta(x) \frac{\partial}{\partial \theta} \frac{\partial}{\partial x} f_\theta(x) - \frac{\partial}{\partial \theta} f_\theta(x) \frac{\partial}{\partial x} f_\theta(x)}{[f_\theta(x)]^2} \end{aligned}$$

for all x and θ. Then the first line established the equivalence with (a) and the second line established the equivalence with (b).

5. (a) For each $y > 0$, the cdf of Y is given by

$$\begin{aligned} \mathbb{P}(Y \leq y) &= \mathbb{P}(-y \leq X \leq y) = \mathbb{P}(-y - \theta \leq X - \theta \leq y - \theta) \\ &= \Phi(y - \theta) - \Phi(-y - \theta). \end{aligned}$$

Hence, differentiating w.r.t. y yields the density of Y:

$$f_\theta(y) = \phi(y - \theta) + \phi(-y - \theta) = \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(y-\theta)^2} + e^{-\frac{1}{2}(y+\theta)^2} \right)$$

for $y > 0$, where $\Phi(x)$ and $\phi(x)$ denote the cdf and pdf of the standard normal distribution respectively. Moreover, by symmetry, the density of Y can be written as

$$\begin{aligned} f_\theta(y) &\equiv \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(y+\min\{-\theta,\theta\})^2} + e^{-\frac{1}{2}(y+\max\{-\theta,\theta\})^2} \right) \\ &\equiv \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(y-|\theta|)^2} + e^{-\frac{1}{2}(y+|\theta|)^2} \right) \end{aligned}$$

which solely depends on $|\theta|$.

(b) From part (a) we can restrict $\theta \geq 0$ WLOG (and to ensure the parametric family is identifiable). Consider the partial derivatives of $f_\theta(y)$:

$$\begin{aligned}
\frac{\partial}{\partial y} f_\theta(y) &= \frac{1}{\sqrt{2\pi}}\left(-(y-\theta)e^{-\frac{1}{2}(y-\theta)^2} - (y+\theta)e^{-\frac{1}{2}(y+\theta)^2}\right), \\
\frac{\partial}{\partial \theta} f_\theta(y) &= \frac{1}{\sqrt{2\pi}}\left((y-\theta)e^{-\frac{1}{2}(y-\theta)^2} - (y+\theta)e^{-\frac{1}{2}(y+\theta)^2}\right), \\
\frac{\partial^2}{\partial \theta \partial y} f_\theta(y) &= \frac{1}{\sqrt{2\pi}}\left(e^{-\frac{1}{2}(y-\theta)^2} - (y-\theta)^2 e^{-\frac{1}{2}(y-\theta)^2} - e^{-\frac{1}{2}(y+\theta)^2} + (y+\theta)^2 e^{-\frac{1}{2}(y+\theta)^2}\right) \\
&= \frac{1}{\sqrt{2\pi}}\left(e^{-\frac{1}{2}(y-\theta)^2} - e^{-\frac{1}{2}(y+\theta)^2} + (y+\theta)^2 e^{-\frac{1}{2}(y+\theta)^2} - (y-\theta)^2 e^{-\frac{1}{2}(y-\theta)^2}\right).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\frac{\partial}{\partial y} f_\theta(y) \frac{\partial}{\partial \theta} f_\theta(y) &= \frac{1}{2\pi}\left((y+\theta)^2 e^{-(y+\theta)^2} - (y-\theta)^2 e^{-(y-\theta)^2}\right), \\
f_\theta(y) \frac{\partial^2}{\partial \theta \partial y} f_\theta(y) &= \frac{1}{2\pi}\big(e^{-(y-\theta)^2} - e^{-(y+\theta)^2} + [(y+\theta)^2 - (y-\theta)^2]e^{-\frac{1}{2}[(y+\theta)^2 + (y-\theta)^2]} \\
&\quad + (y+\theta)^2 e^{-(y+\theta)^2} - (y-\theta)^2 e^{-(y-\theta)^2}\big),
\end{aligned}$$

and

$$\begin{aligned}
f_\theta(y)\frac{\partial^2}{\partial \theta \partial y} f_\theta(y) - \frac{\partial}{\partial y} f_\theta(y) \frac{\partial}{\partial \theta} f_\theta(y) &= \frac{1}{2\pi}\left(e^{-(y-\theta)^2} - e^{-(y+\theta)^2} + 4y\theta e^{-(y^2+\theta^2)}\right) \\
&= \frac{e^{-(y+\theta)^2}}{2\pi}\left(e^{4y\theta} + 4y\theta e^{2y\theta} - 1\right) \geq 0
\end{aligned}$$

by the fact that $y, \theta \geq 0$. Thus, by the results derived in Question 4, we conclude that the density has an MLR in y.

(c) Under the settings in part (b) and by the theorem on Page 35 of Lecture Note 5, we know that $\psi(Y) = \mathbb{1}\{Y > c\}$ is the UMP test based on Y. In order for the test to achieve size α under H_0, we need

$$\alpha = \mathbb{P}_0(Y > c) = \mathbb{P}_0(X > c) + \mathbb{P}_0(X < -c) = 2(1 - \mathbb{P}_0(X \leq c))$$

so $c = z(1 - \frac{\alpha}{2})$, the $(1 - \frac{\alpha}{2})$-th quantile of the standard normal pdf.

(d) The power of $\mathbb{1}\{Y > z(1 - \frac{\alpha}{2})\}$ is

$$\begin{aligned}
\beta_\psi(\theta) &= \mathbb{P}_\theta\left(X - \theta > z\left(1 - \frac{\alpha}{2}\right) - \theta\right) + \mathbb{P}_\theta\left(X - \theta < -z\left(1 - \frac{\alpha}{2}\right) - \theta\right) \\
&= \Phi\left(z\left(\frac{\alpha}{2}\right) + \theta\right) + \Phi\left(z\left(\frac{\alpha}{2}\right) - \theta\right),
\end{aligned}$$

where we used the fact that $-z(p) = z(1-p)$ for $0 \leq p \leq 1$.

Then, we construct another test $\tilde{\psi}(X) = \mathbb{1}\{X < z(\alpha)\}$. It is a level α test since $\mathbb{P}_0(X < z(\alpha)) = \alpha$ under H_0. Moreover, $\tilde{\psi}$ has power

$$\beta_{\tilde{\psi}}(\theta) = \mathbb{P}_\theta(X - \theta < z(\alpha) - \theta) = \Phi(z(\alpha) - \theta).$$

If we consider $H_0 : \theta = 0$ against the simple $H_1 : \theta = -1$ instead, then from Neyman-Pearson Lemma,

$$\begin{aligned}
\mathbb{1}\left\{\frac{f_1(x)}{f_0(x)} = e^{-\frac{2x+1}{2}} > c\right\} &= \mathbb{1}\left\{x < -\log c - \frac{1}{2}\right\} \\
&:= \mathbb{1}\{x < z(\alpha)\} = \tilde{\psi}(x)
\end{aligned}$$

is the UMP test of level α, so it follows that $\beta_{\tilde{\psi}}(-1) \geq \beta_\psi(-1)$ for each α. In particular, when $\alpha = 0.05$, we have $\beta_{\tilde{\psi}}(-1) - \beta_\psi(-1) = 0.0894 > 0$.

```
alpha = 0.05
pnorm(qnorm(alpha)+1)-pnorm(qnorm(alpha/2)-1)-pnorm(qnorm(alpha/2)+1)

## [1] 0.08943598
```

6. (a) Let $0 < \theta_1 < \theta_2$. The likelihood ratio

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{\theta_2}{\theta_1}\frac{\theta_1^2 + x^2}{\theta_2^2 + x^2}, \qquad x \in \mathbb{R}$$

equals to $\theta_1/\theta_2 < 1$ at $x = 0$, but will converge to $\theta_2/\theta_1 > 1$ as $x \to \pm\infty$. Hence it is not monotone in x.

(b) Sufficiency follows from Factorization Theorem with $h(x) = 1$ and $g(|x|\,|\theta) = f(|x|\,|\theta)$.

Using Divide-and-Conquer, the pdf of $|X|$ can be obtained as

$$f_{|X|}(y|\theta) = \frac{2\theta}{\pi}\frac{1}{\theta^2 + y^2}$$

for $y \geq 0$. For $0 < \theta_1 < \theta_2$, the likelihood ratio of $|X|$,

$$\frac{f_{|X|}(y|\theta_2)}{f_{|X|}(y|\theta_1)} = \frac{\theta_2}{\theta_1}\frac{\theta_1^2 + y^2}{\theta_2^2 + y^2}, \qquad y \geq 0,$$

has derivative

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{f_{|X|}(y|\theta_2)}{f_{|X|}(y|\theta_1)} = \frac{\theta_2}{\theta_1}\frac{2y(\theta_2^2 - \theta_1^2)}{[\theta_2^2 + y^2]^2} \geq 0$$

for all $y \geq 0$, so the likelihood ratio is non-decreasing, i.e. the distribution of $|X|$ has MLR.

7. Notice that μ_1 is known. The following analysis focuses on the case that $\mu_1 > 0$. The case for $\mu_1 < 0$ can established symmetrically (i.e. start with the test $\psi(\mathbf{X}) = \mathbb{1}\{\bar{X} \leq c\}$).

(a) Consider the test $\psi(\mathbf{X}) = \mathbb{1}\{\bar{X} \geq c\}$, where $\bar{X} \sim \mathrm{N}(\mu, 1/n)$ and c is a number to be determined. Now, the Type-I and Type-II errors of the test are given by, respectively,

$$\begin{aligned}
\alpha &= \mathbb{P}_{\mu=0}(\bar{X} \geq c) = 1 - \mathbb{P}_{\mu=0}(\sqrt{n}\bar{X} \leq c\sqrt{n}) = 1 - \Phi(c\sqrt{n}) = \Phi(-c\sqrt{n}) \\
1 - \beta &= \mathbb{P}_{\mu=\mu_1}(\bar{X} < c) = \mathbb{P}_{\mu=0}(\sqrt{n}(\bar{X} - \mu_1) \leq (c - \mu_1)\sqrt{n}) = \Phi((c - \mu_1)\sqrt{n}).
\end{aligned}$$

By solving $\alpha = 1 - \beta$, we obtain $c = \mu_1/2$, i.e. we conclude that the test $\psi(\mathbf{X}) = \mathbb{1}\{\bar{X} \geq \mu_1/2\}$ can achieve $\alpha = 1 - \beta$.

(b) In order for the errors to be controlled at level γ, we need

$$\gamma \geq 1 - \Phi(\mu_1\sqrt{n}/2) \iff n \geq \left[\frac{2\Phi^{-1}(1-\gamma)}{\mu_1}\right]^2.$$

(c) Notice that the pdf of X_1 under H_0 and H_1 are, respectively,

$$f_0(x) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}, \quad f_1(x) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_1)^2}{2}\right\},$$

indicating that the likelihood ratio of the first n entries of $\mathbf{X} = (X_1, X_2, \cdots)$ is

$$\lambda_n \equiv \lambda_n(\mathbf{X}) = \prod_{i=1}^{n} \frac{f_1(X_i)}{f_0(X_i)} = \exp\left\{\mu_1 \sum_{i=1}^{n} X_i - \frac{n\mu_1^2}{2}\right\}.$$

In order for the errors to be controlled at level $\alpha = 1 - \beta = \gamma$, we need the thresholds $c_0 = \frac{\gamma}{1-\gamma}$ and $c_1 = \frac{1-\gamma}{\gamma}$, i.e. the SPRT is given as

$$\psi(\mathbf{X}) = \begin{cases} 1, & \text{if } \lambda_\tau \geq \frac{1-\gamma}{\gamma}, \\ 0, & \text{if } \lambda_\tau \leq \frac{\gamma}{1-\gamma} \end{cases},$$

where $\tau := \inf\{n : \lambda_n \geq \frac{1-\gamma}{\gamma} \text{ or } \lambda_n \leq \frac{\gamma}{1-\gamma}\}$ is the time that sample collection was terminated.

Now, since we have

$$\begin{aligned} \mathrm{KL}(f_1||f_0) &= \mathbb{E}_1\left[\log\left(\frac{f_1(X_1)}{f_0(X_1)}\right)\right] = \mathbb{E}_1\left[\mu_1 X_1 - \frac{\mu_1^2}{2}\right] = \frac{\mu_1^2}{2}, \\ -\mathrm{KL}(f_0||f_1) &= \mathbb{E}_0\left[\log\left(\frac{f_1(X_1)}{f_0(X_1)}\right)\right] = \mathbb{E}_0\left[\mu_1 X_1 - \frac{\mu_1^2}{2}\right] = -\frac{\mu_1^2}{2}, \end{aligned}$$

by Wald's Identity,

$$\begin{aligned} \mathbb{E}_1[\tau]\,\mathrm{KL}(f_1||f_0) &= \mathbb{E}_1\left[\log(\lambda_\tau)\right] \\ &\approx \log\left(\frac{1-\gamma}{\gamma}\right)\mathbb{P}_1\left(\lambda_\tau \geq \frac{1-\gamma}{\gamma}\right) + \log\left(\frac{\gamma}{1-\gamma}\right)\mathbb{P}_1\left(\lambda_\tau \leq \frac{\gamma}{1-\gamma}\right) \\ &= (1-\gamma)\log\left(\frac{1-\gamma}{\gamma}\right) + \gamma\log\left(\frac{\gamma}{1-\gamma}\right) = (1-2\gamma)\log\left(\frac{1-\gamma}{\gamma}\right), \end{aligned}$$

indicating that

$$\mathbb{E}_1[\tau] \approx \frac{2(1-2\gamma)\log\left(\frac{1-\gamma}{\gamma}\right)}{\mu_1^2}.$$

Similarly,

$$\begin{aligned} \mathbb{E}_0[\tau](-\mathrm{KL}(f_0||f_1)) &= \mathbb{E}_0\left[\log(\lambda_\tau)\right] \\ &\approx \log\left(\frac{1-\gamma}{\gamma}\right)\mathbb{P}_0\left(\lambda_\tau \geq \frac{1-\gamma}{\gamma}\right) + \log\left(\frac{\gamma}{1-\gamma}\right)\mathbb{P}_0\left(\lambda_\tau \leq \frac{\gamma}{1-\gamma}\right) \\ &= \gamma\log\left(\frac{1-\gamma}{\gamma}\right) + (1-\gamma)\log\left(\frac{\gamma}{1-\gamma}\right) = (2\gamma-1)\log\left(\frac{1-\gamma}{\gamma}\right), \end{aligned}$$

so we also have

$$\mathbb{E}_0[\tau] \approx \frac{2(1-2\gamma)\log\left(\frac{1-\gamma}{\gamma}\right)}{\mu_1^2}.$$

The R.H.S. is the approximated expected number of samples to be collected.

(d) It can be observed graphically that $2[\Phi^{-1}(1-\gamma)]^2 \geq (1-2\gamma)\log\left(\frac{1-\gamma}{\gamma}\right)$ for all $0 < \gamma < 1$, with equality holds iff $\gamma = 0.5$. Hence, the sequential test consumes less samples on average, comparing with the fixed-size test with the same level.