**Mastery.** Let $X_1, X_2, \ldots$ be i.i.d. samples whose distribution depends on an unknown parameter $\theta \in \Theta$. We wish to test the hypotheses $H_0 : \theta = \theta_0$ v.s. $H_1 : \theta = \theta_1$. Let $f_0(x) = f_{X|\Theta}(x|\theta_0)$ and $f_1(x) = f_{X|\Theta}(x|\theta_1)$. For each fixed $m$ let $\lambda_m = \frac{\prod_{i=1}^m f_{X|\Theta}(x_i|\theta_1)}{\prod_{i=1}^m f_{X|\Theta}(x_i|\theta_0)} = \prod_{i=1}^m \frac{f_1(x_i)}{f_0(x_i)}$ denote the likelihood ratio obtained for the first $m$ observations.

The sequential probability ratio test follows the following algorithm, for some thresholds $0 < B < A < \infty$: Start with $m = 1$. If $\lambda_m \leq B$ then the decision is made to accept $H_0$. If $\lambda_m \geq A$ then the decision is made to accept $H_1$. If $B < \lambda_m < A$ then increase the number of observations $m$ by 1.

Let $\alpha_1$ and $\alpha_2$ denote the probabilities of a type I and type II error, respectively. Note that in the paper by Wald, $\beta$ was used instead to denote the probability of a type II error. Let $N$ denote the number of observations needed before a decision is made, with observed value $N = n$. You may assume without proof that $\mathrm{P}(1 \leq N < \infty) = 1$. Let $S = (X_1, \ldots, X_N)$ denote the sequence of observations carried out.

(a) State at least one advantage of the sequential probability ratio test over the standard likelihood ratio test with fixed sample size.

(b) Wald showed that

$$A \leq \frac{1-\alpha_2}{\alpha_1}, \tag{1}$$

$$B \geq \frac{\alpha_2}{1-\alpha_1}. \tag{2}$$

Let $A = 4$ and $B = \frac{1}{4}$. Draw a graph with axes $\alpha_1$, $\alpha_2$ in which the region of probabilities $(\alpha_1, \alpha_2)$ which satisfy the inequalities (1)&(2) is shown as a shaded polygon. Label each vertex of the shaded polygon with its coordinates.

(c) Suppose $X_k \sim \text{Bernoulli}(\theta)$. The sequence of values taken by the $X_k$ is

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | $\cdots$ |
|---|---|---|---|---|---|---|---|---|
| $X_k$ | 1 | 0 | 0 | 1 | 1 | 1 | 0 | $\cdots$ |

Perform a sequential probability ratio test for $H_0 : \theta = 0.5$ v.s. $H_1 : \theta = 0.8$, using thresholds $A = 3$, $B = \frac{1}{3}$.

(d) (i) Let $\mathrm{P}_0$ denote the probability obtained when $\theta = \theta_0$, and similarly let $\mathrm{P}_1$ denote the probability obtained when $\theta = \theta_1$. Explain carefully why, if $(x_1, \ldots, x_n)$ is a sequence of observations which leads to the decision "accept $H_1$", then

$$\mathrm{P}_1(S = (x_1, \ldots, x_n)) \geq A \cdot \mathrm{P}_0(S = (x_1, \ldots, x_n)).$$

*[Hint: $S = (x_1, \ldots, x_n)$ implies that $N = n$.*
*For $i = 0, 1$, $\mathrm{P}_i(S = (x_1, \ldots, x_n) \mid N = n) = \prod_{m=1}^n f_i(x_m)$.]*

(ii) Using (d)(i) explain carefully why $1 - \alpha_2 \geq A\alpha_1$. You may assume for simplicity that the range of $X_k$ is $\mathbb{N}$.
*[Hint: Apply the law of total probability to $1 - \alpha_2$.]*