1. (a) The likelihood function for $Y_1, \ldots, Y_n$ is given as
\[
(2\pi\sigma^2)^{-n/2} \prod_{i=1}^{n} \frac{1}{i^{1/2}} \exp \left\{-\frac{\sum_{i=1}^{n} iY_i^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} iY_i}{\sigma^2} \mu - \frac{\sum_{i=1}^{n} i}{2\sigma^2} \mu^2 \right\}.
\]
This is a 2-parameter exponential family with complete sufficient statistics
\[
\left( \sum_{i=1}^{n} iY_i, \sum_{i=1}^{n} iY_i^2 \right).
\]
(b) Since $E[\sum_{i=1}^{n} iY_i] = \mu \sum_{i=1}^{n} i$, the UMVUE for $\mu$ is
\[
\frac{2\sum_{i=1}^{n} iY_i}{n(n+1)}.
\]
(c) The variance of the UMVUE is
\[
\frac{2\sigma^2}{n(n+1)}.
\]
While, the Fisher information matrix for $\mu$ and $\sigma^2$ is calculated as
\[
\begin{pmatrix}
\sum_{i=1}^{n} i/\sigma^2 & 0 \\
0 & n/(2\sigma^4)
\end{pmatrix}.
\]
The CR-lower bound for $\mu$ is $\sigma^2 / \sum_{i=1}^{n} i$. Thus, the UMVUE attains this bound.

2. (a) We need find $a_1, \ldots, a_n$ minimizing
\[
\text{Var}(\sum_{i=1}^{n} a_iY_i) = \sigma^2 \sum_{i=1}^{n} a_i^2 / i
\]
under constraints $\sum_{i=1}^{n} a_i = 1$ and $a_i > 0, i = 1, \ldots, n$. Using the Lagrange-multiplier or the inequality
\[
\left( \sum_{i=1}^{n} a_i^2 / i \right) \left( \sum_{i=1}^{n} i \right) \geq (\sum_{i=1}^{n} a_i)^2 = 1,
\]
we can obtain that $a_i = i / \sum_{i=1}^{n} i = 2i / (n(n+1))$. That is,
\[
\bar{Y}_n^* = \frac{\sum_{i=1}^{n} iY_i}{n(n+1)/2}.
\]
(b) Note
\[
(\bar{Y}_n^* - \mu) = \frac{\sum_{i=1}^{n} i(Y_i - \mu)}{n(n+1)/2} = \frac{\sum_{i=1}^{n} \sqrt{i} \epsilon_i}{n(n+1)/2}.
\]
Apply the Lindeberg CLT, we obtain
\[
\sqrt{\frac{n(n+1)}{2}}(\bar{Y}_n^* - \mu) \rightarrow_d N(0, \sigma^2).
\]
To verify the Lindeberg condition, note that the total variance $\sigma_n^2 = \sigma^2$ and for any $\delta > 0$,
\[
\sum_{i=1}^{n} E \left[ \frac{i}{n(n+1)/2} \epsilon_i^2 I(\sqrt{i} |\epsilon_i| > \delta \sigma_n \sqrt{\frac{n(n+1)}{2}}) \right] \leq E[\epsilon_i^2 I(\epsilon_i > \sqrt{\frac{n+1}{2}} \sigma_n)] \rightarrow 0.
\]
3. (a) The information matrix for $\mu$ and $\sigma$ is

$$I(\theta) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{pmatrix}.$$  

Since $\dot{q}_\theta = (-\sigma/\mu^2, 1/\mu)$ and $\dot{l}_\theta = (Y - \mu)/\sigma^2, -1/\sigma + (Y - \mu)^2/\sigma^3$, we obtain that the efficient influence function for CV is

$$\dot{q}_\theta I(\theta)^{-1} \dot{l}_\theta = \frac{(Y - \mu)^2 - \sigma^2}{2\mu \sigma} - \frac{\sigma(Y - \mu)}{\mu^2}$$

and that the information bound is

$$\dot{q}_\theta I(\theta)^{-1} \dot{q}_\theta' = \frac{\sigma^4}{\mu^4} + \frac{\sigma^2}{2\mu^2}.$$  

(b) Clearly,

$$\hat{CV} = g(\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2/n, \bar{Y}_n),$$

where $g(x, y) = \sqrt{x}/y$. Since

$$\sqrt{n} \left\{ \left( \sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2/n, \bar{Y}_n \right) - (\sigma^2, \mu) \right\} \rightarrow_d N \left( 0, \begin{pmatrix} 2\sigma^4 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right),$$

the Delta method gives

$$\sqrt{n}(\hat{CV} - \sigma/\mu) \rightarrow_d N(0, \sigma^4/\mu^4 + \sigma^2/(2\mu^2)).$$

(c) We need to maximize the log-likelihood function

$$-n \log \sqrt{2\pi \sigma^2} - \frac{1}{2\sigma^2} \sum_{i=1}^{n}(Y_i - \mu)^2$$

in the region $\mu \geq 1$ and $\sigma > 0$. For fixed $\sigma$, this is a concave quadratic function with mode at $\bar{Y}_n$; thus, it is easy to see that $\hat{\mu}_n = \bar{Y}_n I(\bar{Y}_n \geq 1) + I(\bar{Y}_n < 1)$. Then $\hat{\sigma}_n = \sqrt{\sum_{i=1}^{n}(Y_i - \hat{\mu}_n)^2/n}$. Note

$$\sqrt{n}(\bar{Y}_n - \hat{\mu}_n) = \sqrt{n}(\bar{Y}_n - 1)I(\bar{Y}_n < 1)$$

and for $\epsilon > 0$,

$$P(\sqrt{n}(\bar{Y}_n - 1)I(\bar{Y}_n < 1) \geq \epsilon) \leq P(\bar{Y}_n < 1) \rightarrow_p 0$$

from the assumption that the true mean is larger than 1. We then conclude that $\sqrt{n}(\sum_{i=1}^{n}Y_i^2/n, \bar{Y}_n)$ should have the same asymptotic distribution as $\sqrt{n}(\sum_{i=1}^{n}Y_i^2/n, \hat{\mu}_n)$. Therefore, $\hat{\sigma}_n/\hat{\mu}_n$ has the same asymptotic distribution as $\hat{CV}$. One may also sought the MLE theory to obtain the distribution.
4. (a) The observed likelihood function is

\[
\prod_{i=1}^{n} \left\{ \phi(Y_i - \mu) \right\}^{R_i} \left\{ \Phi(y_0 - \mu) \right\}^{1-R_i},
\]

where \( \phi \) and \( \Phi \) are respectively the density and CDF of the standard normal distribution.

(b) For \( n = 1 \), if \( \mu_1 \) and \( \mu_2 \) give the same likelihood function, consider \( R = 1 \), i.e., \( Y > y_0 \), we obtain

\[
\phi(Y - \mu_1) = \phi(Y - \mu_2).
\]

This should be true for any \( Y > y_0 \). It is easy to conclude \( \mu_1 = \mu_2 \).

(c) The likelihood equation for \( \mu \) is

\[
\sum_{i=1}^{n} \left[ (Y_i - \mu)R_i - (1 - R_i) \frac{\phi(y_0 - \mu)}{\Phi(y_0 - \mu)} \right] = 0.
\]

The Newton-Raphson iteration is

\[
\mu^{(k+1)} = \mu^{(k)} - \left\{ -\sum_{i=1}^{n} R_i + \sum_{i=1}^{n} (1 - R_i) \left[ \frac{\phi'(y_0 - \mu^{(k)})}{\Phi(y_0 - \mu^{(k)})} - \frac{\phi(y_0 - \mu^{(k)})^2}{\Phi(y_0 - \mu^{(k)})^2} \right] \right\}^{-1} \sum_{i=1}^{n} \left[ (Y_i - \mu^{(k)})R_i - (1 - R_i) \frac{\phi(y_0 - \mu^{(k)})}{\Phi(y_0 - \mu^{(k)})} \right].
\]

(d) The complete score equation for \( \mu \) is

\[
\sum_{i=1}^{n} (Y_i - \mu) = 0.
\]

Thus, in the M-step, we have

\[
\mu^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} E[Y_i | R_1, Y_1, ..., R_n, Y_n, R_1, ..., R_n; \mu^{(k)}] = \frac{1}{n} \left\{ \sum_{i=1}^{n} (R_i Y_i + (1 - R_i) E[Y_i | Y_i > y_0; \mu^{(k)}]) \right\}.
\]

Clearly, the E-step evaluates \( E[Y_i | Y_i > y_0; \mu^{(k)}] \) using

\[
\int_{y_0}^{\infty} y \phi(y - \mu^{(k)}) dy \frac{1}{1 - \Phi(y_0 - \mu^{(k)})} = \mu^{(k)} + \frac{\phi(y_0 - \mu^{(k)})}{1 - \Phi(y_0 - \mu^{(k)})}.
\]

(e) By the MLE theory, \( \sqrt{n}(\hat{\mu}_n - \mu) \to_d N(0, I(\mu)^{-1}) \), where

\[
I(\mu) = E \left[ R - (1 - R) \left\{ \frac{\phi'(y_0 - \mu)}{\Phi(y_0 - \mu)} - \frac{\phi(y_0 - \mu)^2}{\Phi(y_0 - \mu)^2} \right\} \right]
\]

\[
= 1 - \Phi(y_0) - \phi'(y_0 - \mu) + \frac{\phi(y_0 - \mu)^2}{\Phi(y_0 - \mu)}.
\]

(f) The confidence interval can be constructed as

\[
(\hat{\mu}_n - \Phi^{-1}(0.975) \sqrt{I(\hat{\mu}_n)^{-1}}, \hat{\mu}_n + \Phi^{-1}(0.975) \sqrt{I(\hat{\mu}_n)^{-1}}).
\]