Suppose that \( x = (y_1, y_2, \ldots, y_n) \) has iid normal components with mean \( 0 \) and variance \( V \). Then
\[
f(y|V) = \frac{1}{\sqrt{2\pi V}} \exp\left(-\frac{y^2}{2V}\right)
\]
\[
\ln f(y|V) = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln V - \frac{y^2}{2V}
\]
\[
\frac{\partial}{\partial V} \ln f(y|V) = -\frac{1}{2V} + \frac{y^2}{2V^2}
\]
\[
\frac{\partial^2}{\partial V^2} \ln f(y|V) = +\frac{1}{2V^2} - \frac{y^2}{V^3}
\]

So \( I_y(V) = -E_V \left( \frac{\partial^2}{\partial V^2} \ln f(y|V) \right) = -\frac{1}{2V^2} + \frac{V}{V^3} = \frac{1}{2V^2} \)

To find an MLE for \( V \), we note that
\[
\ln f(x|V) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln V - \frac{1}{2V} \sum y_i^2
\]
and \( \frac{\partial}{\partial V} \ln f(x|V) = -\frac{n}{2V} + \frac{1}{V^2} \sum y_i^2 \)

The likelihood equation \( \frac{\partial}{\partial V} \ln f(x|V) = 0 \) has the unique root
\[
V = \frac{1}{n} \sum y_i^2
\]
which also maximizes the likelihood. That is,
\[
\hat{V}_n^{MLE} = \frac{1}{n} \sum y_i^2
\]

Maximum likelihood theory implies that \( \hat{V}_n^{MLE} \) is approximately normal with mean \( V \) and variance
\[
\frac{1}{nI_y(V)} = \frac{1}{n \left( \frac{1}{2V^2} \right)} = \frac{2V^2}{n}
\]

This produces valid but unusable large \( n \) approximate confidence limits for \( V \).
\( \hat{\theta}_{\text{MLE}} \pm z \sqrt{\frac{2}{n}} \)

This can be "fixed" in the usual ways. The "expected information" fix replaces \( n \ln(\theta) \) with \( n \ln(\hat{\theta}_{\text{MLE}}) \) ultimately producing

\[
\hat{\theta}_{\text{MLE}} = \left( 1 \pm z \sqrt{\frac{2}{n}} \right)
\]

The "observed information" fix replaces \( n \ln(\theta) \) with the negative curvature of the log-likelihood at the MLE.

\[
- \frac{d^2}{d\theta^2} \ln f(x | \theta) = -\frac{n}{2\theta^2} + \frac{1}{2\theta^3} \sum y_i^2
\]

When this is evaluated at \( \theta = \frac{1}{n} \sum y_i^2 \) we get the same formula as for \( \theta_{\text{MLE}} \).

---

Here's an interesting complement to the above (not presented in class). In this example, earlier material can be used to get different, but nearly equivalent intervals for \( \theta \). Definition 5 implies that as it is a sum of independent squared standard normal variables,

\[
\sum \left( \frac{x_i^2}{V} \right)^2 = \frac{\sum x_i^2}{V} \sim \chi^2_n
\]

So

\[
P \left[ \frac{1}{n} \text{ (lower 2.5\%) } < \frac{\hat{\theta}_{\text{MLE}}}{V_n} < \frac{1}{n} \text{ (upper 2.5\%) } \right] = 0.5
\]

i.e.

\[
P \left[ \frac{\hat{\theta}_{\text{MLE}}}{V_n} < V < \frac{\hat{\theta}_{\text{MLE}}}{n \chi^2_n} \right] = 0.5
\]

But then, since for \( z_c \) i.i.d \( N(0, 1) \)

\[
\sum z_c^2 \sim \chi^2_n
\]
and the CLT says that

\[ \frac{1}{n} \sum \frac{Z_i^2}{\mu_i^2} \text{ is approximately } N \left( 1, \frac{2}{n} \right) \]

or \[ \sum \frac{Z_i^2}{\mu_i^2} \text{ is approximately } N(n, 2n) \]. So

upper 2.5% pt of \( \chi_n^2 \) \( \approx n + 1.96 \sqrt{2n} \)

town 2.5% pt of \( \chi_n^2 \) \( \approx n - 1.96 \sqrt{2n} \)

So \( \chi \) says that approximate 95% confidence limits for \( \mu \) are

\[ \frac{\hat{\mu}_{MLE}}{\frac{1}{n} (n+1.96 \sqrt{2n})} \quad \text{and} \quad \frac{\hat{\mu}_{MLE}}{\frac{1}{n} (n-1.96 \sqrt{2n})} \]

i.e.

\[ \frac{\hat{\mu}_{MLE}}{1 + 1.96 \sqrt{2/n}} \quad \text{and} \quad \frac{\hat{\mu}_{MLE}}{1 - 1.96 \sqrt{2/n}} \]

Then, since for small \( |x| \)

\[ \frac{1}{1+x} \approx 1 - x \]

these are approximately the same the limits derived from maximum likelihood theory.