function of σ^2, is $1/(2\sigma^4)$. If we use this expression for the information matrix, evaluated at $\hat{\sigma}^2$, the Wald test becomes

$$W_1 = \frac{n}{2}\hat{\sigma}^{-4}\big(\hat{\sigma}^2 - 1\big)^2 = \frac{n}{2}e^{-4\hat{\tau}}\big(e^{2\hat{\tau}} - 1\big)^2 = 2n\hat{\tau}^2 + o(1).$$

Since this differs from (13.57), we have shown that different parametrizations lead to numerically different Wald statistics even if the true information matrix, evaluated at the MLE of the model parameter, is used in both cases.

As we will see in the next section, the LM test is invariant if it is based on the true information matrix evaluated at the MLE. But if some other information matrix estimator is used, the LM test can also be parametrization dependent. Suppose that we use the empirical Hessian. From (13.48), the first two derivatives of ℓ with respect to σ^2, evaluated at $\sigma^2 = 1$, are

$$\begin{aligned} \frac{\partial \ell}{\partial \sigma^2}\bigg|_{\sigma^2=1} &= -\frac{1}{2}\bigg(n - \sum_{t=1}^{n} y_t^2\bigg) = \frac{n}{2}\big(e^{2\hat{\tau}} - 1\big) \quad \text{and} \\ \frac{\partial^2 \ell}{(\partial \sigma^2)^2}\bigg|_{\sigma^2=1} &= \frac{n}{2}\big(1 - 2e^{2\hat{\tau}}\big). \end{aligned}$$

From this, we find that the statistic LM_2 calculated as was (13.56) but for the σ^2 parametrization, is

$$LM_2 = \frac{n\big(e^{2\hat{\tau}} - 1\big)^2}{2\big(2e^{2\hat{\tau}} - 1\big)} = 2n\hat{\tau}^2 + o(1). \tag{13.58}$$

The leading term is correct, as it must be, but (13.58) is numerically different from (13.56).

Plainly, there are still more forms of both the LM and Wald tests, some but not all of which will coincide with one of the versions we have already computed. The interested reader is invited to try out, for example, the effects of using σ itself, rather than σ^2, as the model parameter.

This example illustrates the fact that there may be many different classical tests, which are numerically different but asymptotically equivalent. The fact that there are so many different tests creates the problem of how to choose among them. One would prefer to use tests that are easy to compute and for which the finite-sample distribution is well approximated by the asymptotic distribution. Unfortunately, it frequently requires considerable effort to determine the finite-sample properties of asymptotic tests. Any method of analysis tends to be restricted to very special cases, such as the case of linear regression models with normal errors discussed in Section 13.4. One generally applicable approach is to use computer simulation (Monte Carlo experiments); see Chapter 21.