from which we may derive that

$$e^{2\hat{\tau}} = \hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} y_t^2. \tag{13.50}$$

For this parametrization, the information matrix, which has only one element, is constant and equal to 2:

$$\mathfrak{I} = -\frac{1}{n}E(D_\tau^2 \ell) = \frac{2}{n}\sum_{t=1}^{n} e^{-2\tau}E(y_t^2) = 2.$$

Notice that, although $\mathfrak{I}$ is constant, the loglikelihood function is *not* a quadratic function of τ. We now consider various classical tests for the null hypothesis that $\tau = 0$, or, equivalently, that $\sigma^2 = 1$. Despite the simplicity of this example, we will uncover a bewildering variety of test statistics.

Initially, we will work with the τ parametrization. It is not necessary to do any estimation at all in order to find restricted estimates, since $\tilde{\tau} = 0$. For the Wald and LR tests we need to find $\hat{\tau}$. From (13.50), it is

$$\hat{\tau} = \frac{1}{2}\log\left(\frac{1}{n}\sum_{t=1}^{n} y_t^2\right).$$

The restricted "maximum" of the loglikelihood function is just the value of the function at $\tau = 0$:

$$\tilde{\ell} = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n} y_t^2 = -\frac{n}{2}\log 2\pi - \frac{n}{2}e^{2\hat{\tau}}. \tag{13.51}$$

Although this is the restricted maximum, it is convenient to express it, as we have done here, in terms of the unrestricted estimate, $\hat{\tau}$. The unrestricted maximum, $\hat{\ell}$, is given by

$$-\frac{n}{2}\log 2\pi - n\hat{\tau} - \frac{1}{2}e^{-2\hat{\tau}}\sum_{t=1}^{n} y_t^2 = -\frac{n}{2}\log 2\pi - n\hat{\tau} - \frac{n}{2}, \tag{13.52}$$

where the equality uses (13.50).

We may proceed at once to obtain the LR statistic, which is twice the difference between (13.52) and (13.51):

$$\begin{aligned} LR = 2(\hat{\ell} - \tilde{\ell}) &= n\big(e^{2\hat{\tau}} - 1 - 2\hat{\tau}\big) \\ &= 2n\hat{\tau}^2 + o(1). \end{aligned} \tag{13.53}$$

The second line of (13.53) is a Taylor expansion of the statistic in powers of $\hat{\tau}$. This is of interest because, *under the null hypothesis*, we expect $\hat{\tau}$, which is both the estimate itself and the *difference* between the estimate and the true value of the parameter, to be of order $n^{-1/2}$. It follows that $2n\hat{\tau}^2$ will be of order unity and that higher terms in the expansion of the exponential function in (13.53) will be of lower order. Thus, if the various forms of the classical test do indeed yield asymptotically equal expressions, we may expect that the leading term of all of them will be $2n\hat{\tau}^2$.

Let us next consider the LM statistic. The essential piece of it is the derivative of the loglikelihood function (13.49) with respect to τ, evaluated at $\tau = 0$. We find that

$$\frac{\partial \ell}{\partial \tau} = -n + e^{-2\tau} \sum_{t=1}^{n} y_t^2 \quad \text{and} \quad \frac{\partial \ell}{\partial \tau}\bigg|_{\tau=0} = n\big(e^{2\hat{\tau}} - 1\big). \tag{13.54}$$

If for the variance of $\partial \ell / \partial \tau$ we use n times the true, constant, value of the single element of the information matrix, 2, the LM statistic is the square of $(\partial \ell / \partial \tau)|_{\tau=0}$, given by (13.54), divided by $2n$:

$$LM_1 = \frac{n}{2}\big(e^{2\hat{\tau}} - 1\big)^2 = 2n\hat{\tau}^2 + o(1).$$

This variant of the LM statistic has the same leading term as the LR statistic (13.53) but will of course differ from it in finite samples.

Instead of the true information matrix, an investigator might prefer to use the negative of the empirical Hessian to estimate the information matrix; see equations (8.47) and (8.49). Because the loglikelihood function is not exactly quadratic, this estimator does *not* coincide numerically with the true value. Since

$$\frac{\partial^2 \ell}{\partial \tau^2} = -2e^{-2\tau} \sum_{t=1}^{n} y_t^2, \tag{13.55}$$

which at $\tau = 0$ is $-2ne^{2\hat{\tau}}$, the LM test calculated in this fashion is

$$LM_2 = \frac{n}{2} e^{-2\hat{\tau}} \big(e^{2\hat{\tau}} - 1\big)^2 = 2n\hat{\tau}^2 + o(1). \tag{13.56}$$

The leading term is as in LR and LM_1, but LM_2 will differ from both those statistics in finite samples.

Another possibility is to use the OPG estimator of the information matrix; see equations (8.48) and (8.50). This estimator is

$$\frac{1}{n} \sum_{t=1}^{n} \left(\frac{\partial \ell}{\partial \tau}\right)^2 = \frac{1}{n} \sum_{t=1}^{n} \big(y_t^2 e^{-2\tau} - 1\big)^2,$$

which, when evaluated at $\tau = 0$, is equal to

$$\frac{1}{n} \sum_{t=1}^{n} \big(y_t^2 - 1\big)^2.$$