This expression cannot even be expressed as a function of $\hat{\tau}$ alone. To obtain an expansion of the test statistic that makes use of it, we must make use of the property of the normal distribution which tells us that $E(y_t^4) = 3\sigma^4$, or, in terms of τ, $3e^{4\tau}$.[4] Using this property, we can invoke a law of large numbers and conclude that the OPG information matrix estimator is indeed equal to $2 + o(1)$ at $\tau = 0$. Thus the third variant of the LM test statistic is

$$LM_3 = \frac{n^2\big(e^{2\hat{\tau}} - 1\big)^2}{\sum_{t=1}^n \big(y_t^2 - 1\big)^2} = 2n\hat{\tau}^2 + o(1).$$

Once again, the leading term is $2n\hat{\tau}^2$, but the form of LM_3 is otherwise quite different from that of LM_1 or LM_2.

Just as there are various forms of the LM test, so are there various forms of the Wald test. Any one of these may be formed by combining the unrestricted estimate $\hat{\tau}$ with some estimate of the information matrix, which in this case is actually a scalar. The simplest choice is just the true information matrix, that is, 2. With this we obtain

$$W_1 = 2n\hat{\tau}^2. \tag{13.57}$$

It is easy to see that W_2, which uses the empirical Hessian, is identical to W_1, because (13.55) evaluated at $\tau = \hat{\tau}$ is just $-2n$. On the other hand, use of the OPG estimator yields

$$W_3 = \hat{\tau}^2 \sum_{t=1}^{n} \big(y_t^2 e^{-2\hat{\tau}} - 1\big)^2,$$

which is quite different from W_1 and W_2.

All of the above test statistics were based on τ as the single parameter of the model, but we could just as well use σ or σ^2 as the model parameter. Ideally, we would like test statistics to be invariant to such reparametrizations. The LR statistic is always invariant, since $\hat{\ell}$ and $\tilde{\ell}$ do not change when the model is reparametrized. But all forms of the Wald statistic, and some forms of the LM statistic, are in general not invariant, as we now illustrate.

Suppose we take σ^2 to be the parameter of the model. The information matrix is not constant in this new parametrization, and so we must evaluate it at the *estimate* $\hat{\sigma}^2$. It is easy to see that the information matrix, as a

[4] Note that it was *not* necessary to use special properties of the normal distribution in order to expand the previous statistics, which were in fact all functions of one and only one random variable, namely $\hat{\tau}$. In general, in less simple situations, this agreeable feature of the present example is absent and special properties must be invoked in order to discover the behavior of all the various test statistics.