the standard normal distribution, this probability is

$$P(\alpha, \lambda) \equiv 1 - \Phi(c_\alpha - \lambda) + \Phi(-c_\alpha - \lambda). \tag{12.36}$$

In order to find the inverse power function corresponding to (12.36), we let $P(\alpha, \lambda) = \pi$ for some desired level of power π. This equation implicitly defines the inverse power function. It is easy to check from (12.36) that $P(\alpha, -\lambda) = P(\alpha, \lambda)$. Thus, if $P(\alpha, \lambda) = \pi$, then $P(\alpha, -\lambda) = \pi$ also. However, the nonuniqueness of λ would not arise if we were to square the test statistic to obtain a χ^2 form. No closed-form expression exists giving the (absolute) value of λ as a function of α and π in the present example, but for any given arguments λ is not hard to calculate numerically.

What interpretation should we give to the resulting function $\lambda(\alpha, \pi)$? If we square the asymptotically normal statistic (12.35) in order to obtain a χ^2 form, the result will have a limiting distribution of $\chi^2(1, \Lambda)$ with $\Lambda = \lambda^2$. Then it appears that $\Lambda = (\lambda(\alpha, \pi))^2$ is asymptotically the smallest NCP needed in order that a test of size α based on the square of (12.35) should have probability at least π of rejecting the null.

Let the nonlinear regression model be written, as usual, as

$$\boldsymbol{y} = \boldsymbol{x}(\boldsymbol{\beta}) + \boldsymbol{u}, \tag{12.37}$$

where the parameter of interest θ is a component of the parameter vector $\boldsymbol{\beta}$. If we denote by $\boldsymbol{X}_\theta$ the derivative of the vector $\boldsymbol{x}(\boldsymbol{\beta})$ with respect to θ, evaluated at the parameters $\boldsymbol{\beta}_0$, and by $\boldsymbol{M}_X$ the projection off all the columns of $\boldsymbol{X}(\boldsymbol{\beta})$ other than $\boldsymbol{X}_\theta$, then the asymptotic variance of the least squares estimator $\hat{\theta}$ is $\sigma_0^2(\boldsymbol{X}_\theta^\top \boldsymbol{M}_X \boldsymbol{X}_\theta)^{-1}$, where σ_0^2 is the variance of the components of $\boldsymbol{u}$. If we consider a DGP with a parameter $\theta \neq \theta_0$, then for a given sample size n, the parameter δ of the drifting DGP becomes $n^{1/2}(\theta - \theta_0)$, and $\Lambda = \lambda^2$ becomes

$$\Lambda = \frac{1}{\sigma_0^2}(\theta - \theta_0)^2 \boldsymbol{X}_\theta^\top \boldsymbol{M}_X \boldsymbol{X}_\theta. \tag{12.38}$$

This may be compared with the general expression (12.26). Now let $\theta(\alpha, \pi)$ be the value of θ that makes Λ in (12.38) equal to $(\lambda(\alpha, \pi))^2$ as given above by the inverse power function. We see that, within an asymptotic approximation, DGPs with values of θ closer to the θ_0 of the null hypothesis than $\theta(\alpha, \pi)$ will have probability less than π of rejecting the null on a test of size α.

We should be unwilling to regard a failure to reject the null as evidence against some other DGP or set of DGPs if, under the latter, there is not a fair probability of rejecting the null. What do we mean by a "fair probability" here? Some intuition on this matter can be obtained by considering what we would learn in the present context by using a standard tool of conventional statistical inference, namely, a confidence interval. Armed with the estimate $\hat{\theta}$ and an estimate of its standard error, $\hat{\sigma}_\theta$, we can form a confidence interval