Transformation of parameters: hypothesis tests

- The three large-sample tests does not behave the same under reparameterization (transformation of parameters)
- The Wald test is not invariant to parameter transformation
- The score test and the likelihood ratio tests are invariant to transformation of parameter

Hypothesis test

Suppose we wish to test $H_0: \theta = \theta_0$ for scalar parameter θ.

The Wald test statistic is $Z = \sqrt{I(\hat{\theta})}(\hat{\theta} - \theta_0)$.

An approximate test with α level of significance is defined by the rejection rule

$$\text{Reject } H_0 \text{ if } |Z| \geq z_{\alpha/2}.$$

Now suppose $\phi = \Phi(\theta)$ is an equivalent parameterization of the model. Then $H_0: \theta = \theta_0$ can be expressed equivalently as

$$H_0: \phi = \phi_0$$

where $\phi_0 = \Phi(\theta_0)$.

The corresponding Wald test statistic is $Z = \sqrt{I(\hat{\phi})}(\hat{\phi} - \phi_0)$.

Invariance of hypothesis tests

For independent Bernoulli observations $y_1, y_2, \ldots, y_n$ the hypothesis $H_0\!: \theta = 0.5$ can be expressed equivalently as $H_0\!: \phi = 0$ if

$$\phi = \log\left(\frac{\theta}{1-\theta}\right).$$

However, it can be checked that the Wald test statistics corresponding to the two equivalent formulations of the problem are not equal.

The fact that the conclusion of the Wald test can be influenced by the choice of parameterization is a serious criticism of the approach.

If a test gives identical conclusions irrespective of the parametrization, it is said to be *invariant*.

The Wald test is not invariant, but the score test and log-likelihood ratio test are.

The invariance property of the score test will be given in Theorem 16.

Theorem 16

Suppose $y_1, y_2, \ldots, y_n$ are independent observations with log-likelihood function

$$\ell_\theta(\theta; \boldsymbol{y}) = \ell_\phi(\phi; \boldsymbol{y})$$

where $\phi = \Phi(\theta)$. Consider the hypothesis

$$H_0: \theta = \theta_0 \iff H_0: \phi = \phi_0$$

and let u_θ and u_ϕ be the score statistics defined from the two log-likelihood functions.

If Φ is $1-1$ and onto and twice continuously differentiable with $\Phi'(\theta) \neq 0$ then

$$\boxed{|u_\phi| = |u_\theta|.}$$

Proof of Theorem 16

Notations	$\log L$	score	Fisher information
θ form	$\ell_\theta(\theta;y)$	$S_\theta(\theta;y)$	I_θ
ϕ form	$\ell_\phi(\phi;y)$	$S_\phi(\phi;y)$	I_ϕ

Recall that $\ell_\phi(\Phi(\theta);y) = \ell_\theta(\theta;y)$

(apply chain rule)

$$S_\theta(\theta;y) = \frac{\partial}{\partial\theta}\ell_\theta(\theta;y) = \frac{\partial}{\partial\theta}\ell_\phi(\Phi(\theta);y) = \underbrace{\frac{\partial \ell_\phi}{\partial\phi}\frac{\partial\phi}{\partial\theta}} = S_\phi(\phi;y)\,\Phi'(\theta)$$

$$\frac{\partial^2\ell_\theta}{\partial\theta} = \frac{\partial}{\partial\theta}\left(\boxed{\frac{\partial\ell_\theta}{\partial\theta}}\right) = \frac{\partial}{\partial\theta}\left[\frac{\partial\ell_\phi}{\partial\phi}\frac{\partial\phi}{\partial\theta}\right] = \frac{\partial^2\ell_\phi}{\partial\phi^2}\left(\frac{\partial\phi}{\partial\theta}\right)^2 + \left(\frac{\partial\ell_\phi}{\partial\phi}\right)\left(\frac{\partial^2\phi}{\partial\theta^2}\right)$$

(apply product rule)

$$I_\theta = E\left[-\frac{\partial^2\ell_\theta}{\partial\theta}\right]$$

$$= E\left[-\left(\frac{\partial^2\boxed{\ell_\phi}}{\partial\phi^2}\right)\left(\frac{\partial\phi}{\partial\theta}\right)^2 - \left(\frac{\partial\boxed{\ell_\phi}}{\partial\phi}\right)\left(\frac{\partial^2\phi}{\partial\theta^2}\right)\right]$$

$$= \left(\frac{\partial\phi}{\partial\theta}\right)^2 E\left[-\frac{\partial^2\ell_\phi}{\partial\phi^2}\right] - \left(\frac{\partial^2\phi}{\partial\theta^2}\right)E\left[\frac{\partial\ell_\phi}{\partial\phi}\right]$$

$$= \left(\frac{\partial\phi}{\partial\theta}\right)^2 I_\phi - \left(\frac{\partial^2\phi}{\partial\theta^2}\right)\underbrace{E[S_\phi]}_{=0}$$

under regularity conditions

Proof of Theorem 16

$$I_\theta = \left(\frac{\partial \phi}{\partial \theta}\right)^2 I_\phi = \Phi'(\theta)^2 I_\phi$$

$$U_\theta = \frac{S_\theta(\theta_0; y)}{\sqrt{I_{\theta_0}}}$$

$$= \frac{S_\phi(\phi_0; y)\, \Phi'(\theta_0)}{\sqrt{\Phi'(\theta_0)^2 I_{\phi_0}}}$$

$$= \frac{S_\phi(\phi_0; y)\, \cancel{\Phi'(\theta_0)}}{\sqrt{I_{\phi_0}} \quad \pm \cancel{\Phi'(\theta_0)}}$$

$$= \pm \frac{S_\phi(\phi_0; y)}{\sqrt{I_{\phi_0}}}$$

$$= \pm U_\phi$$

$$\Rightarrow \quad |U_\theta| = |U_\phi| .$$

Example 5.14

Suppose $y_1, y_2, \dots, y_n$ are $i.i.d.\, Po(\lambda)$ observations. Let $\phi = \log(\lambda)$. Find the score test statistic $\mathcal{U}_\phi$.

Recall $\ell(\lambda; y) = \left(\sum_{i=1}^{n} y_i\right)\log\lambda - n\lambda - \log\left(\prod_{i=1}^{n} y_i!\right)$

$$\ell(\phi; y) = \left(\sum_{i=1}^{n} y_i\right)\log(e^\phi) - ne^\phi - \log\left(\prod_{i=1}^{n} y_i!\right)$$

$\phi = \log\lambda \Rightarrow \lambda = e^\phi$

$$= \left(\sum_{i=1}^{n} y_i\right)\phi - ne^\phi - \log\left(\prod_{i=1}^{n} y_i!\right)$$

$$S(\phi; y) = \left(\sum_{i=1}^{n} y_i\right) - ne^\phi$$

$$\frac{\partial^2 \ell_\phi}{\partial\phi^2} = -ne^\phi$$

$$I_\phi = E\left[-\frac{\partial^2 \ell_\phi}{\partial\phi^2}\right] = E\left[ne^\phi\right] = ne^\phi$$

$$\mathcal{U}_\phi = \frac{S_\phi(\phi_0; y)}{\sqrt{I_{\phi_0}}} = \frac{\sum_{i=1}^{n} y_i - ne^{\phi_0}}{\sqrt{ne^{\phi_0}}} = \frac{n\bar{y} - ne^{\phi_0}}{\sqrt{ne^{\phi_0}}} = \frac{\bar{y} - e^{\phi_0}}{\sqrt{\frac{e^{\phi_0}}{n}}} = \boxed{\frac{\bar{y} - \lambda_0}{\sqrt{\frac{\lambda_0}{n}}}}$$

$\phi_0 = \log(\lambda_0)$

same as Example 5.11

Invariance property of the likelihood ratio test

It is straightforward to check that the likelihood ratio test is invariant under transformations of the parameter θ.

$$G^2_\theta = -2\left[\ell_\theta(\theta_0;y) - \ell_\theta(\hat{\theta};y)\right]$$

$$= -2\left[\ell_\phi(\Phi(\theta_0);y) - \ell_\phi(\Phi(\hat{\theta});y)\right]$$

$$= -2\left[\ell_\phi(\phi_0;y) - \ell_\phi(\hat{\phi};y)\right]$$

$$= G^2_\phi$$

So G^2 is invariant under transformation of θ.

There is no simple relationship between G^2 and $\mathcal{U}$.
But for large samples, we have $G^2 \approx \mathcal{U}$.