Confidence Intervals

- Theorems 13 and 14 gave the distributional results about the score and the maximum likelihood estimator (MLE)
- We can use the these as pivotal quantities to build confidence intervals and hypothesis tests for θ
- We will build three (large-sample) tests: the Wald test, the score test, and the likelihood ratio test
- They are based on the MLE, the score, and the ratio of the likelihood at θ_0 and $\hat{\theta}$, respectively.

Approximate confidence intervals

Suppose $y_1, y_2, \dots, y_n$ are independent observations with log-likelihood $\ell(\theta; \boldsymbol{y})$.

An approximate $100(1-\alpha)\%$ confidence interval for θ is given by

$$\boxed{\left(\hat{\theta} - z_{\alpha/2}\sqrt{I_{\hat{\theta}}^{-1}}, \hat{\theta} + z_{\alpha/2}\sqrt{I_{\hat{\theta}}^{-1}}\right).}$$

Theorem 14: $\hat{\theta} \longrightarrow N(\theta, I_{\theta}^{-1})$

Use this as a pivotal quantity to build a CI for θ.

$CI = \hat{\theta} \pm Z_{\frac{\alpha}{2}} \sqrt{I_{\hat{\theta}}^{-1}}$

θ is unknown. Use $\hat{\theta}$ as an approximation.

The hypothesis test constructed using $\hat{\theta}$ above as the pivotal quantity is called the Wald test. The above CI is called as Wald confidence interval.

Wald test statistic

Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. observations with log-likelihood $\ell(\theta; \boldsymbol{Y})$.

The Wald test statistic for

$$H_0: \theta = \theta_0 \qquad \text{vs} \quad H_a: \theta \neq \theta_0$$

is given by

$$\boxed{Z = \frac{\hat{\theta} - \theta_0}{\sqrt{I_{\hat{\theta}}^{-1}}}.}$$

If $H_0: \theta = \theta_0$ is true, then the distribution of Z converges to $N(0,1)$ as $n \to \infty$.

A test with significance level approximately α is given by the rule:

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

Score test statistic

Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. observations with log-likelihood $\ell(\theta; \boldsymbol{Y})$.

The score test statistic for

$$H_0: \theta = \theta_0$$

is given by

$$\boxed{U = \frac{S(\theta_0; \boldsymbol{Y})}{\sqrt{I_{\theta_0}}}.}$$

We can also use Theorem 13 to construct a hypothesis test for θ. This test is known as the Score test.

If $H_0: \theta = \theta_0$ is true, then the distribution of U converges to $N(0,1)$ as $n \to \infty$.

A test with significance level approximately α is given by the rule:

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

if doing a two-sided test

$H_a: \theta \neq \theta_a$

Log-likelihood ratio test statistic

Suppose $Y_1, Y_2, \dots, Y_n$ are i.i.d. observations with log-likelihood $\ell(\theta; \boldsymbol{Y})$.

The log-likelihood ratio test statistic for

$$H_0: \theta = \theta_0$$

is given by

$$\boxed{G^2 = -2\left[\ell(\theta_0; \boldsymbol{Y}) - \ell(\hat{\theta}; \boldsymbol{Y})\right].}$$

$$G^2 = -2 \log \frac{L(\theta_0; y)}{L(\hat{\theta}; y)} = -2 \log L(\theta_0; y) + 2 \log L(\hat{\theta}; y) = -2\,\ell(\theta_0; y) + 2\,\ell(\hat{\theta}; y)$$

If $H_0: \theta = \theta_0$ is true then, under suitable regularity conditions, the distribution of G^2 converges to $\boxed{\chi^2_1}$ as $n \to \infty$.

A test with significance level approximately α is given by the rule:

$$\text{Reject } H_0 \text{ if } \boxed{G^2 \geq \chi^2_{1,\alpha}}.$$

Example 5.11

Suppose $y_1, y_2, \dots, y_n$ are $i.i.d.$ $\boxed{Po(\lambda)}$ observations.
Suppose we wish to test $\boxed{H_0: \lambda = \lambda_0.}$
Calculate the Wald, Score, and log-likelihood ratio test statistics.

Recall $\hat{\lambda} = \bar{y}$, $S(\lambda;\theta) = -n + \frac{1}{\lambda}\left(\sum_{i=1}^{n} y_i\right)$, $I_\lambda = \frac{n}{\lambda}$

Wald: $Z = \dfrac{\hat{\theta} - \theta_0}{\sqrt{I_{\hat{\theta}}^{-1}}} = \dfrac{\hat{\lambda} - \lambda_0}{\sqrt{I_{\hat{\lambda}}^{-1}}} = \dfrac{\bar{y} - \lambda_0}{\sqrt{\frac{\hat{\lambda}}{n}}} = \dfrac{\bar{y} - \lambda_0}{\sqrt{\frac{\bar{y}}{n}}}$

Score: $U = \dfrac{S(\lambda_0; y)}{\sqrt{I_{\lambda_0}}} = \dfrac{-n + \frac{1}{\lambda_0}\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\frac{n}{\lambda_0}}} = \dfrac{-n + \frac{n\bar{y}}{\lambda_0}}{\sqrt{\frac{n}{\lambda_0}}} = \dfrac{\left(\frac{n}{\lambda_0}\right)(-\lambda_0 + \bar{y})}{\sqrt{\frac{n}{\lambda_0}}}$

$= \dfrac{\bar{y} - \lambda_0}{\sqrt{\frac{\lambda_0}{n}}}$

Example 5.11 Solution

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

$$G^2 = -2\left[\ell(\lambda_0; y) - \ell(\hat{\lambda}; y)\right]$$

$$= -2\left[-n\lambda_0 + \left(\sum_{i=1}^{n} y_i\right)\log\lambda_0 + \cancel{\log\prod_{i=1}^{n}\left(\frac{1}{y_i!}\right)} - \left(-n\hat{\lambda} + \left(\sum_{i=1}^{n} y_i\right)\log\hat{\lambda} + \cancel{\log\prod_{i=1}^{n}\left(\frac{1}{y_i!}\right)}\right)\right]$$

$$= -2\left[-n(\lambda_0 - \hat{\lambda}) + \left(\sum_{i=1}^{n} y_i\right)(\log\lambda_0 - \log\hat{\lambda})\right]$$

$$= -2\left[-n(\lambda_0 - \bar{y}) + n\bar{y}\log\left(\frac{\lambda_0}{\hat{\lambda}}\right)\right]$$

$$= -2n\left[(\lambda_0 - \bar{y}) + \bar{y}\log\left(\frac{\lambda_0}{\bar{y}}\right)\right]$$

Exercise: Write down all three tests formally (i.e. state the hypotheses, test statistic, and critical region/P-value).

Example 5.11 Solution

Exercise: Give an approximate $100(1-\alpha)\%$ confidence interval for λ.

We will give the Wald confidence interval:

$$\mathrm{CI} = \hat{\lambda} \pm z_{\alpha/2}\sqrt{I_{\lambda}^{-1}}$$

$$= \bar{y} \pm z_{\alpha/2}\sqrt{\frac{\bar{y}}{n}}$$