Lecture 27: Interval estimation

Ciaran Evans

Motivation

Suppose we have data $(X_1, Y_1), ..., (X_n, Y_n)$ with

$$Y_i \sim Bernoulli(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta^T X_i$$

So far, we have discussed:

- Finding point estimates $\widehat{\beta}$
- Testing hypotheses about the true (but unknown) parameters β

Question: What are the limitations of point estimates and hypothesis tests for inference about β?

Confidence interval

```
##                 Estimate   Std. Error   z value       Pr(
## (Intercept)  2.641506279 0.1213233066  21.77246 4.23334
## WBC         -0.289290446 0.0134349261 -21.53272 7.68928
## PLT         -0.006561464 0.0005932064 -11.06101  1.9389
```

Question: How would I calculate a 95% confidence interval for β_1 (the change in the log odds of dengue for a one-unit increase in WBC, holding PLT fixed)?

Confidence interval

```
##                 Estimate   Std. Error   z value       Pr(
## (Intercept)  2.641506279 0.1213233066  21.77246 4.233346
## WBC         -0.289290446 0.0134349261 -21.53272 7.689284
## PLT         -0.006561464 0.0005932064 -11.06101  1.93894
```

95% confidence interval for β_1: (-0.315, -0.262)

Question: How do I interpret this confidence interval?

Deriving the coverage probability

$$(1-\alpha) \text{ Wald interval} : \widehat{\theta} \pm z_{\alpha/2} SE(\widehat{\theta})$$

Formal definition

Inverting a test

Theorem: Let $\theta \in \Theta$ be a parameter of interest. For each value of $\theta_0 \in \Theta$, consider testing $H_0 : \theta = \theta_0$ vs. $H_A : \theta \neq \theta_0$, and let $\mathcal{R}(\theta_0)$ be the rejection region for a level α test. Let $C(X_1, ..., X_n) = \{\theta_0 \in \Theta : (X_1, ..., X_n) \notin \mathcal{R}(\theta_0)\}$. Then $C(X_1, ..., X_n)$ is a $1 - \alpha$ confidence set for θ.

Example

Suppose $X_1, ..., X_n \overset{iid}{\sim} \mathit{Uniform}[0, \theta]$. Let's invert the LRT to find a confidence interval for θ.