Interval estimation

Motivation

Suppose we have data $(X_1, Y_1), \dots, (X_n, Y_n)$ 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

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

Confidence interval

```
...
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.6415063  0.1213233   21.77   <2e-16 ***
## WBC         -0.2892904  0.0134349  -21.53   <2e-16 ***
## PLT         -0.0065615  0.0005932  -11.06   <2e-16 ***
## ---
...
```

> 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(>|z|)
## (Intercept)  2.6415063  0.1213233   21.77   <2e-16 ***
## WBC         -0.2892904  0.0134349  -21.53   <2e-16 ***
## PLT         -0.0065615  0.0005932  -11.06   <2e-16 ***
## ---
...
```

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

> How do I interpret this confidence interval?

Deriving the coverage probability

Formal definition

Inverting a test

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform[0, \theta]$. We want to test

$$H_0 : \theta = \theta_0 \qquad H_A : \theta \neq \theta_0$$

> Find the LRT statistic for this test.

Example

Suppose $X_1, \ldots, X_n \stackrel{iid}{\sim} Uniform[0, \theta]$. Inverting the LRT gives us a confidence interval of the form

$$C(X_1, \ldots, X_n) = \{\theta : X_{(n)} \leq \theta \leq X_{(n)}k'\}$$

> Find a value k' such that the test is size α.