

Linear Regression

Outcome: Continuous

e.g. Birth weight in grams

Covariates: Continuous, categorical, etc.

$$X_1 = \begin{cases} 1 & \text{History of pre-term labor} \\ 0 & \text{No history of PTL} \end{cases}$$

$$X_2 = \text{age (years)}$$

Link Function: None

- Model Birth weight directly

"

Model:

~~Blanks~~

- Expected value of birthweight given covariates

$$E[y | X_1, X_2] = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

↑ "given"
"Expected value"

General Interpretation

β_0 : Intercept. The expected value of birth weight when $X_1=0$ and $X_2=0$. In this model, $X_2=0$ means age=0, so β_0 has no practical interpretation.

β_1 : Slope. Controlling for age, β_1 is the change in birthweight comparing women with a Hx of PTL ($X_1=1$) to women without a Hx of PTL ($X_1=0$)

Logistic Regression

(1)

Outcome: Binary

$$\text{Low} = \begin{cases} 1 & \text{If birth weight} \leq 2500\text{g} \\ 0 & \text{If birth weight} > 2500\text{g} \end{cases}$$

Covariates: Continuous, categorical, etc.

$$X_1 = \begin{cases} 1 & \text{Hx of PTL} \\ 0 & \text{No Hx of PTL} \end{cases}$$

$$X_2 = \text{age (years)}$$

Link Function: logit

$$\text{logit}(\pi) = \log\left(\frac{\pi}{1-\pi}\right)$$

"log odds"

Model:

- Probability of low birth weight given covariates

- Uses logit link function

$$\text{logit}(\Pr(\text{Low}=1 \text{ given covariates})) = \beta_0^* + \beta_1^* X_1 + \beta_2^* X_2$$

General Interpretation

β_0^* : Intercept. The direct interpretation of β_0^* lacks a practical meaning.

β_0^* is the logit of the probability of having a low birth weight baby when $X_1=0$ and $X_2=0$. To get a practical interpretation, we need to switch from the "logit of the probability..." to the actual "probability of low birth weight given $X_1=0$ "

β_0 : slope. Controlling for a history of PTL, β_0 is the change in birth weight associated with a 1 year increase in age.

To do this switch, we need to use the inverse of the logit function. If $XB = \beta_0^* + x_1\beta_1^* + x_2\beta_2^*$,

$$\log\left(\frac{\pi}{1-\pi}\right) = XB$$

- Solving for π using algebra,

$$\pi = \frac{1}{1 + \exp(-XB)}$$

- We can interpret $\frac{1}{1 + \exp(-\beta_0^*)}$ as the prob. of low birthweight when $x_1=0$ and $x_2=0$.

β_1^* : Controlling for age, β_1^* is the log odds ratio for low birth weight. Comparing women ~~not~~ ^{with} a history of PTL ~~to~~ women with no history of PTL. $e^{\beta_1^*}$ is the odds ratio, and is what we always report

β_2^* : ~~Also~~ a log odds ratio, but we interpret $e^{\beta_2^*}$. $e^{\beta_2^*}$ is the multiplicative change (or a "fold increase") in the odds of having a low birth weight child for a one-year increase in age.

Model Output/Estimates

Linear Regression Model

```
lm(formula = bwt ~ ptl2 + age, data = lbw.data)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2657.728	233.078	11.403	< 2e-16 ***
ptl2	-456.142	142.127	-3.209	0.00157 **
age	15.463	9.828	1.573	0.11733

Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05

$$BWT = 2657.7 - 456.1 * PTL + 15.463 * age$$

Model Output/Estimates

Logistic Regression Model

```
lrm(formula = low ~ ptl2 + age, data = lbw.data)
```

	Coef	S.E.	Wald Z	P
Intercept	0.53225	0.77460	0.69	0.4920
ptl2	1.60766	0.42986	3.74	0.0002
age	-0.07054	0.03411	-2.07	0.0386

$$\text{logit}(\Pr(LBW=1)) =$$

$$0.53 + 1.61 * PTL - 0.07 * age$$

Hypothesis Tests

(1) Association between mean birthweight and pre-term labor

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$T = \frac{-456.142 - 0}{142.127} = -3.209$$

$$p\text{-value} = .00157 \quad (n=3 \text{ d.f.})$$

Hypothesis Tests

(1) Association between low birth

Weight and pre-term labor

$$H_0: \beta_1^* = 0$$

$$H_0: e^{\beta_1^*} = 1$$

$$H_A: \beta_1^* \neq 0$$

$$H_A: e^{\beta_1^*} \neq 1$$

$$Z = \frac{1.60766 - 0}{.42986} = 3.74$$

$$p\text{-value} = .0002 \quad (\text{Normal dist'n; } n=3 \text{ d.f.})$$

(2) Could also test age

(2) Could also test age

Confidence Intervals: (95% CI)

(1) Pre-term Labor

$$-456.142 \pm t_{n-3, 1-\alpha/2} \times 142.127 \quad (n=189)$$

Appendix,

$$[-736.55, -175.75]$$

- History of PTL is associated with a 456.1 gram decrease in birth weight (95% CI: [-736.6, -175.8]), holding age constant.

(2) Could also do a CI for a 1-year increase in age, holding PTL constant.

Confidence Intervals

(1) Pre-term labor

Step 1: Calculate on log odds ratio scale

$$\hat{\beta}_1^* \pm Z_{.975} s(\hat{\beta}_1^*) \\ 1.60766 \pm 1.96(.42896) \text{ & corrector} \\ = [.77, 2.45]$$

Step 2: Convert to odds ratios

$$[e^{.77}, e^{2.45}]$$

$$[2.15, 11.59]$$

- Controlling for age, Pre-term labor is associated with a 4.99 fold increased odds of having a low birth weight child (95% CI = [2.15, 11.59]).

(2) CI for age would follow same two steps (1) obtain CI for log odds ratio, (2) use exp to convert to odds ratio

Prediction

(1) What is the predicted birth weight if the mother has no history of PTL ($X_1=0$) and age is 27 ($X_2=27$)?

- Recall from the output:

$$\text{BWT} = 2657.7 - 456.1(X_1) + 15.463(X_2)$$

- For $X_1=0, X_2=27$

$$\text{BWT} = 3075.2 \text{ grams}$$

Prediction

(1) What is the predicted probability of having a low birth weight child if $X_1=0, X_2=27$?

- From Output

$$\text{logit}(\pi) = 0.532 + 1.61(X_1) - 0.07(X_2)$$

- For $X_1=0, X_2=27$

$$\text{logit}(\pi) = -1.372$$

- We now need to switch to the probability scale.

Recall that if

$$\text{logit}(\pi) = XB$$

Then

$$\pi = \frac{1}{1 + \exp(-XB)}$$

$$- \text{Prob of LBW} = \frac{1}{1 + \exp(-(-1.372))}$$

$$= \frac{1}{1 + \exp(1.372)}$$

$$= 0.2022$$

if $X_1=0, X_2=27$