

4.1 $\text{logit}(\pi) = 5.085 - 0.1156 \cdot \text{temp}$, where $\text{se}(\beta_0) = 3.052$ and $\text{se}(\beta_1) = 0.047$ Standard Error

① If true slope for β_1 is assumed to be 0, what is the probability that you estimate $|\hat{\beta}_1|$ to be 0.1156 or larger?

→ If β_1 is in fact 0, any estimate that we get is due to pure random noise.

Logistic regression estimates are approximately normal, because of the CLT and the properties of MLE

Hence assume $\hat{\beta}_1 \sim N(0, \text{se}^2)$

→ Convert observed values into Z-score: $z = \frac{\hat{\beta}_1 - 0}{\text{se}} = \frac{-0.1156 - 0}{0.047} = -2.46$

Z-score of -2.46 means that 2.46 standard deviations below 0

$P(|Z| > 2.46) = 2 \times P(Z \geq 2.46)$ (since double-sided)

$P(Z \geq 2.46) \approx 0.007$, and $2 \times 0.007 = 0.014$

= If the true slope of β_1 is 0, there is only about 1.4% of seeing an estimate of 0.1156.

② $\hat{\beta}_1 = 0.1156$ means that each additional Fahrenheit temperature decreases the odds of O-Ring failure by $\exp(-0.1156) = 0.891$ ↖ (due to -0.1156)

③ For a decrease of 10 degrees in temperature (Fahrenheit), how do odds of O-Ring failure increase?

$= e^{(-0.1156)(-10)} = 3.18$. O-Ring failure increases by a factor of 3.18.

④ Assume $x=31$. What is $P(Y=\text{Failed} | X)$?

$\log \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$

$\log \frac{p}{1-p} = 5.085 - 0.1156 \times 31$

$\frac{p}{1-p} = e^z$

$p = (1-p)e^z$

$p = e^z - e^z p$

$p + e^z p = e^z$

$p(1 + e^z) = e^z$

$p = \frac{e^z}{1 + e^z} = \frac{e^{1.5014}}{1 + e^{1.5014}} = 0.817...$

⑤ What would the estimated coefficient be if the temperature were measured in Celsius?

Formula = $(X^\circ\text{C} \cdot \frac{9}{5}) + 32 = 1.8^\circ\text{C} + 32$.

$$\text{logit}(p) = 5.085 - 0.1156(1.8 \cdot x + 32)$$

$$= 1.3858 - 0.20808C$$

B_1 = P-values do not change since unit rescales both the estimates and SE

B_0 = Intercept will change

4.2 y = Survival, x = passenger class. Consider $\text{logit}(\text{survived} | \text{class}, i)$ = factor variable

Survived	Yes	No	Yes	No	No	Yes
Class	Crew	Second	First	Third	Crew	Second

	X_{crew}	X_1	X_2	X_3
1	1	0	0	0
2	0	0	1	0
3	0	1	0	0
4	0	0	0	1
5	1	0	1	0
6	0	0	0	0

4.3 Women who gave birth under overcrowding: who did not $\Rightarrow 1.8:1$ odds ratio

$$\frac{\text{Odds}_0}{\text{Odds}_n} = 1.8$$

① $\log(1.8) = 0.588$ is $\hat{B}_1$ $\log\left(\frac{P_n}{1-P_n}\right) = B_0$, $\log\left(\frac{P_0}{1+P_0}\right) = B_0 + B_1x$ where $x=1$

② $P(y = \text{Caesarian section}) = 0.4$ $\log\left(\frac{\frac{P_0}{1+P_0}}{\frac{P_n}{1-P_n}}\right) = B_1$ $\log\left(\frac{\text{Odds}_0}{\text{Odds}_n}\right) = B_1$

$x=1$ as overcrowding

$B_1 = \log(1.8)$

What is the probability of Caesarian if she had not been subject to overcrowding?

$$P(\text{Maria's probability of a Caesarean section}) = 0.4$$

$$\text{Odds of overpowering} = \frac{P}{1-P} = \frac{0.4}{0.6} = \frac{2}{3} = 0.667$$

From ①, we know that $\frac{\text{Odds}_0}{\text{Odds}_n} = 1.8$

Then, $\frac{0.667}{\text{Odds}_n} = 1.8$. $\text{Odds}_n = \frac{0.667}{1.8} = 0.370$

$$\frac{P}{1-P} = 0.370 \quad p = 0.370 - 0.370 \cdot P$$

$$P(1+0.370) = 0.370 \quad P = \frac{0.370}{1+0.370} = 0.270$$

9.4 Main effect = individual effect of a single variable on the outcome, independent of other variables

(*) $\beta_1 x_1 + \beta_2 x_2$, the main effects are just their individual coefficients

Non-linear model = ones with parameters (β) appear nonlinearity, not the predictors

(*) $y = \frac{\beta_1}{1 + e^{-\beta_2 x}}$ β_1 and β_2 cannot be separated.