Interpretation of the Beta coefficients from a Multiple Regression Model

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

where $\beta_0, \beta_1, \ldots, \beta_k$ are unknown parameters that must be estimated.

Interpretation of model parameters

β_0: y-intercept of $(k+1)$-dimensional surface; the value of $E(y)$ when $x_1 = x_2 = \cdots = x_k = 0$

β_1: Change in $E(y)$ for a 1-unit increase in x_1, when $x_2, x_3, \ldots, x_k$ are held fixed

β_2: Change in $E(y)$ for a 1-unit increase in x_2, when $x_1, x_3, \ldots, x_k$ are held fixed

$\vdots$

β_k: Change in $E(y)$ for a 1-unit increase in x_k, when $x_1, x_2, \ldots, x_{k-1}$ are held fixed

Interpretation of the Beta coefficients from a Multiple Logistic Regression Model

Logistic Model

$$\pi = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}$$

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

Logit transform from pi to pi*

$$\pi^* = \ln\left(\frac{\pi}{1-\pi}\right)$$

$$\pi^* = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

Interpretations of β Parameters in the Logistic Model

$$\pi^* = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

where

$$\pi^* = \ln\left(\frac{\pi}{1-\pi}\right)$$

$\pi = P(y = 1)$

β_i = Change in log-odds π^* for every 1-unit increase in x_i, holding all other x's fixed

$e^{\beta_i} - 1$ = Percentage change in odds $\pi/(1-\pi)$ for every 1-unit increase in x_i, holding all other x's fixed

Need to multiply by 100 to obtain the percentage change in odds

```
> #Multiple Logistic Regression in R
> if (FALSE)
+ {"
+ Consider a study of the analgesic effects of treatments on elderly patients with neuralgia.
+ Two test treatments and a placebo are compared. The response variable is whether the patient reported pain or not.
+ Researchers recorded the age and gender of 60 patients and the duration of complaint before the treatment began.
+ Treatment = A and B represent the two test treatments and P represents the placebo treatment
+ Duration = duration of complaint (months) before treatment
+ Pain is the response variable 0=No Pain, 1=Pain
+ "}
```

```
>
> logistic <- glm(Pain ~ Treatment + Sex + Age + Duration, family=binomial(logit),
data=neuralgia)
> summary(logistic)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -20.588282   7.102883  -2.899  0.00375 **
TreatmentB   -0.526853   0.937025  -0.562  0.57394
TreatmentP    3.181690   1.016021   3.132  0.00174 **
SexM          1.832202   0.796206   2.301  0.02138 *
Age           0.262093   0.097012   2.702  0.00690 **
Duration     -0.005859   0.032992  -0.178  0.85905
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P(Y=1)/(1-P(Y=1)) = odds

$$\frac{P(Y=1|Age=x+1)/(1-P(Y=1|Age=x+1))}{P(Y=1|Age=x)/(1-P(Y=1|Age=x))} = \textbf{Age odds ratio = 1.2996}$$

```
>
> #obtain the odds ratios for each coefficient
> exp( coef(logistic))
 (Intercept)   TreatmentB   TreatmentP         SexM          Age     Duration
1.144518e-09 5.904604e-01 2.408742e+01 6.247630e+00 1.299648e+00 9.941584e-01
>
> #obtain the percentage change in odds (pi/(1-pi)) for every
> # 1-unit increase in Xi, holding all other X's fixed

> (exp(coef(logistic)) - 1) * 100
 (Intercept)   TreatmentB   TreatmentP         SexM          Age     Duration
 -99.9999999  -40.9539573 2308.7421759  524.7629579   29.9647809   -0.5841551
>
> ##-----------------------##
>
```