

Case study

Low birth weight data: The goal of this study was to identify risk factors associated with giving birth to a low birth weight baby (weighing less than 2500 grams). Data were collected on 189 women, 59 of which had low birth weight babies and 130 of which had normal birth weight babies. Four variables which were thought to be of importance were age, weight of the subject at her last menstrual period, race, and the number of physician visits during the first trimester of pregnancy.

Variables:

Variable	Abbreviation
Identification Code	ID
Low Birth Weight (0 = Birth Weight \geq 2500g, 1 = Birth Weight < 2500g)	LOW
Age of the Mother in Years	AGE
Weight in Pounds at the Last Menstrual Period	LWT
Race (1 = White, 2 = Black, 3 = Other)	RACE
Smoking Status During Pregnancy (1 = Yes, 0 = No)	SMOKE
History of Premature Labor (0 = None 1 = One, etc.)	PTL
History of Hypertension (1 = Yes, 0 = No)	HT
Presence of Uterine Irritability (1 = Yes, 0 = No)	UI
Number of Physician Visits During the First Trimester (0 = None, 1 = One, 2 = Two, etc.)	FTV
Birth Weight in Grams	BWT

Low birth weight data: Background information

Low birth weight is an outcome that has been of concern to physicians for years. This is due to the fact that infant mortality rates and birth defect rates are very high for low birth weight babies. A woman's behavior during pregnancy (including diet, smoking habits, and receiving prenatal care) can greatly alter the chances of carrying the baby to term and, consequently, of delivering a baby of normal birth weight.

The variables identified in the code sheet given in the table have been shown to be associated with low birth weight in the obstetrical literature. The goal of the current study was to ascertain if these variables were important in the population being served by the medical center where the data were collected.

Use the binary variable LOW as your response variable.

Analysis 1:

- Considering the four most clinically important variables previously discussed, i.e., age, weight of the subject at her last menstrual period, race, and the number of physician visits during the first trimester of pregnancy, carry out a model building exercise to determine the best model to describe the data. Based your model building exercise on the AIC and AIC weights. Discuss your results carefully.
- Use backward elimination procedure to decide which predictor variables can be dropped from the regression model. Control the type I error at $\alpha = .10$ at each stage. Which variables are retained? How does this compare to the results obtained with the AIC? Discuss.

Analysis 2:

- Consider now the 8 covariates included in the data set (all except Birth Weight in Grams). Repeat the previous model building exercise using the AIC and the AIC weights. Compare the results of both analysis and make a careful discussion.
- Use backward elimination procedure to decide which predictor variables can be dropped from the regression model. Control the type I error at $\alpha = .10$ at each stage. Which variables are retained? How does this compare to the results obtained with the AIC? Discuss.

Conclusion:

Using the final model selected in the previous model building exercise (based on the AIC) answer the following scientific questions

- What scientific insight does your final model offer?
- Interpret the final model using the estimated coefficients and odd ratios.
- How confident do you feel about the model you selected? Discuss.