

CAM625 Module 3 Assignment

Interpretation of statistical results

Instructions

Your submission should consist of a word document addressing the questions set out below. The number of words suggested next to each part below is guide for approximately what we expect. They are not strict limits, but you should be brief and your responses should be clear and concise. We expect short answers rather than an essay. This assignment is worth 10% of your final mark.

Please do not discuss (neither personally nor on the discussion board) specific findings or results from this assignment. We anticipate that you may have some queries but please try to phrase your questions carefully when posting to the discussion board. If in doubt contact us both privately by email:

Petr.Otahal@utas.edu.au and Karen.Wills@utas.edu.au

Due Date The Module 3 Assignment is due for submission in the [MyLO Dropbox](#) at 5pm on Sunday 25th September. The Final Deadline for submission is 12pm Thursday 29th September.

Low Birth Weight Data set

The dataset we will be using for this assessment task is the Low Birth Weight data collected at Baystate Medical Center, Springfield, Massachusetts during 1986, and used as an example of fitting a multiple logistic regression model in the text: *Applied Logistic Regression*: Third Edition by Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X. (2013).

Story behind the data

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

List of variables

ID - Identification Code

LOW - Low Birth Weight (0 = Birth Weight \geq 2500g, 1 = Birth Weight < 2500g)

low_f - Factor variable for LOW (No = Birth Weight \geq 2500g, Yes = Birth Weight < 2500g)

AGE - Age of the Mother in Years

LWT - Mother's Weight in Pounds at the Last Menstrual Period

RACE - Race (1 = White, 2 = Black, 3 = Other)

SMOKE - Smoking Status During Pregnancy (0 = No, 1 = Yes)

smoke_f - Factor variable for SMOKE (No/Yes)

PTL - History of Premature Labor (0 = None 1 = One, etc.)

ptl_f - Factor variable for PTL recoded as a binary variable (No/Yes)

HT - History of Hypertension (0 = No, 1 = Yes)

ht_f - Factor variable for HT (No/Yes)

UI - Presence of Uterine Irritability (0 = No, 1 = Yes)

ui_f - Factor variable for UI (No/Yes)

FTV - Number of Physician Visits During the First Trimester, (0 = None, 1 = One, 2 = Two, etc.)

BWT Birth Weight in Grams

Assignment overview

This assignment is about the interpretation of statistical results.

For the assignment we will consider the associations of low birth weight with mother's history of hypertension, smoking status during pregnancy, age and mother's weight at last menstrual period. You will not be required to perform any analysis; the analyses have been done and your task is to interpret the results. The assignment will focus on one key relationship under investigation, however you will also be required to interpret some other associations that may be of interest.

Outcome variable The outcome variable to be used for the Module 3 Learning Activity and Assignment is *BWT*, the continuous variable for birth weight in grams.

Exposure variables The primary exposure of interest is *ht_f*, the factor variable coding for history of hypertension.

Confounders Potential confounders of the association between birth weight and hypertension are:

- *smoke_f*, smoking status during pregnancy
- *AGE*, Mother's age in years,
- *LWT*, Mother's Weight in pounds at the last menstrual period.

Assignment Tasks

1. Summary statistics (100-200 words)

Examine the participant characteristics in Table 1 which are stratified by the binary variable *ht_f*, indicating whether the individual has a history of hypertension. Provide a **brief** summary for each of the variables describing the differences or similarities based on hypertension status. Use only the data provided, do not run any statistical tests.

Table 1: Participant characteristics stratified by history of hypertension

	Total	Hypertension	
		No	Yes
	(N=189)	(N=177)	(N=12)
Age (years)			
Mean (SD)	23.2 (5.30)	23.3 (5.36)	22.9 (4.44)
Smoking status during pregnancy			
No	115 (61 %)	108 (61 %)	7 (58 %)
Yes	74 (39 %)	69 (39 %)	5 (42 %)
History of preterm labour			
No	159 (84 %)	149 (84 %)	10 (83 %)
Yes	30 (16 %)	28 (16 %)	2 (17 %)
Birth weight (grams)			
Mean (SD)	2945 (729)	2972 (709)	2537 (917)
Presence of uterine irritability			
No	161 (85 %)	149 (84 %)	12 (100 %)
Yes	28 (15 %)	28 (16 %)	0 (0 %)
Weight at last menstrual period (pounds)			
Mean (SD)	130 (30.6)	128 (28.4)	158 (47.0)

2. Bivariate relationships (100-200 words)

a) Figures 1 to 3 are graphical displays of the relationships between birth weight (*BWT*) and:

- *ht_f*, history of hypertension
- *smoke_f*, smoking status during pregnancy
- *AGE*, Mother's age in years,
- *LWT*, Mother's Weight in pounds at the last menstrual period. Briefly describe the patterns you observe from the plots.

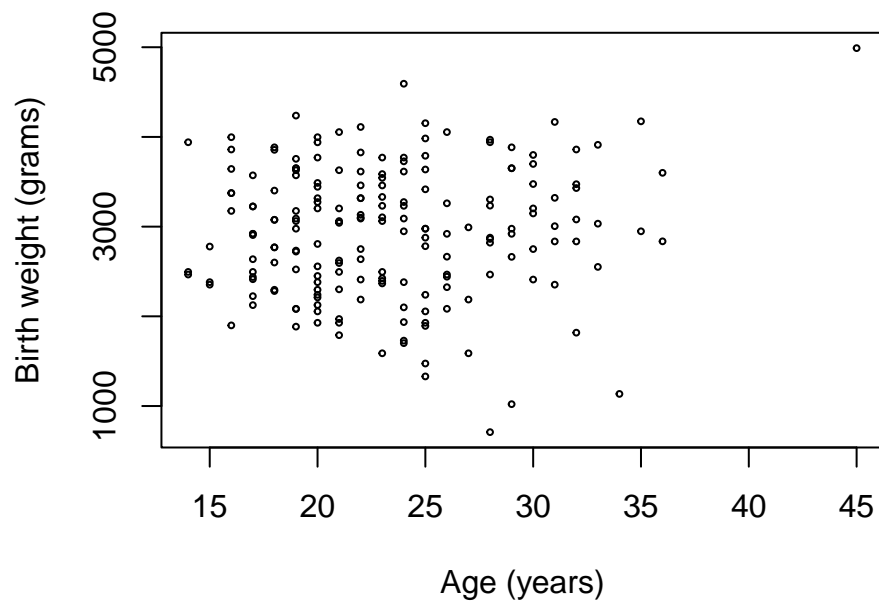


Figure 1: Scatter plot of birth weight and mother's age

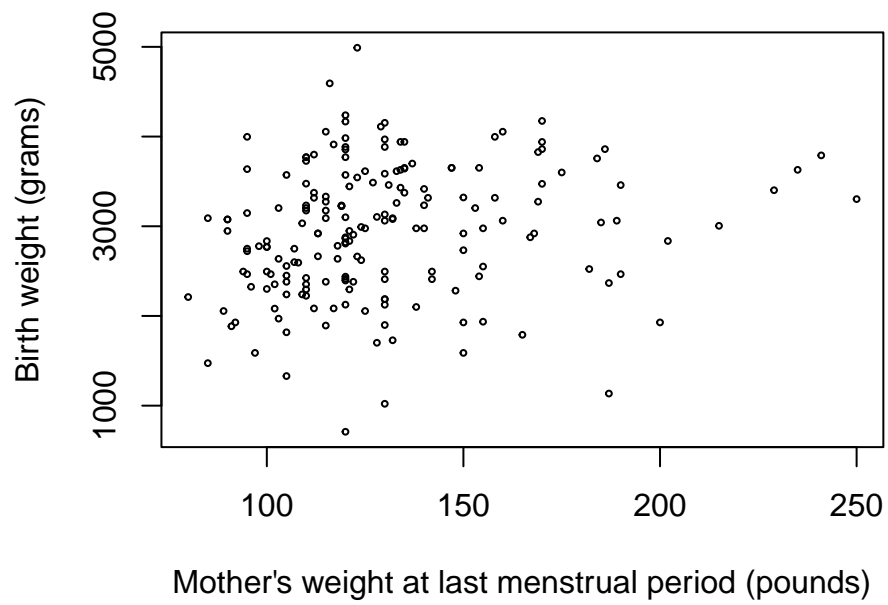


Figure 2: Scatter plot of birth weight and mother's weight

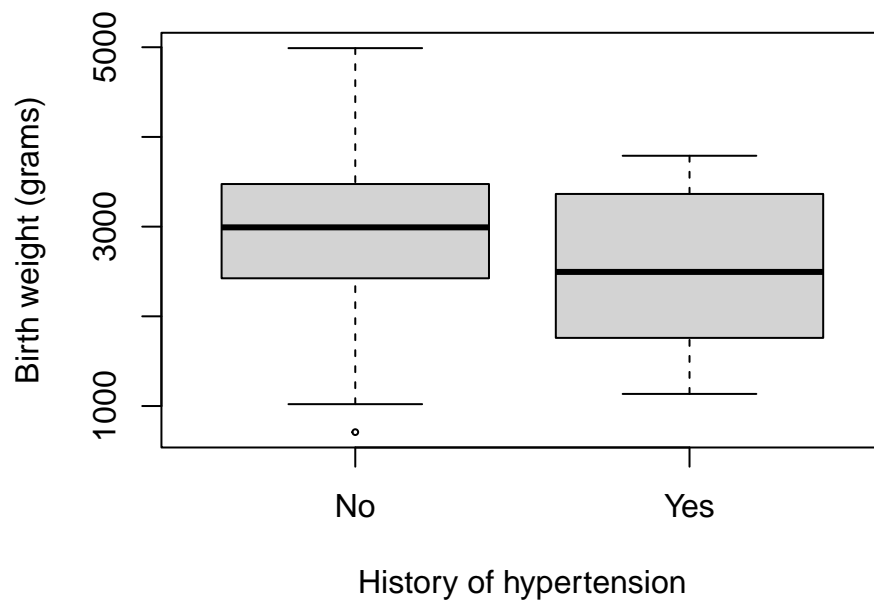


Figure 3: Boxplot for birth weight by history of hypertension

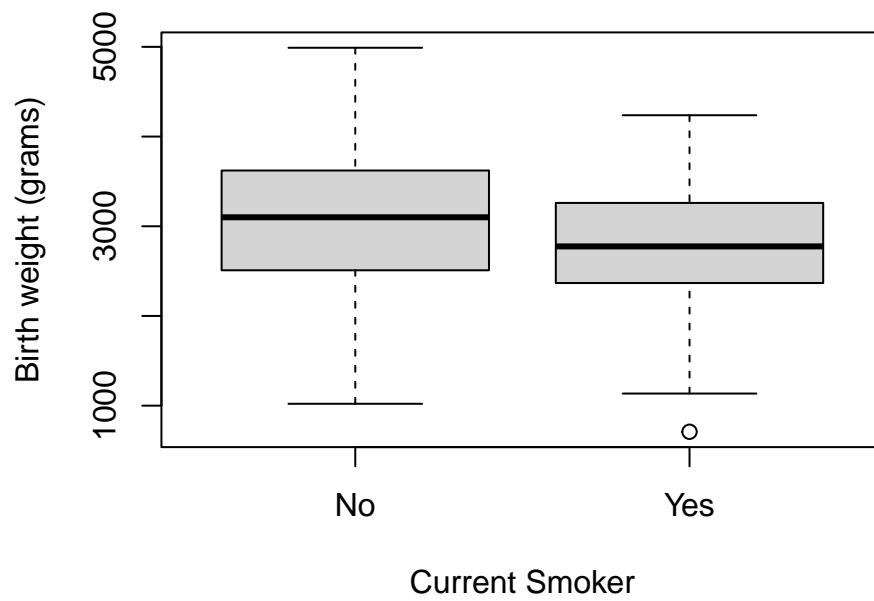


Figure 4: Boxplot for birth weight by smoking status during pregnancy

- b) Table 2 displays the correlation coefficients, 95% confidence intervals and p-values for the association of birth weight with mother's age, *AGE* and weight at last menstrual period, *LWT*. Comment on the magnitude of the association indicated by the test statistic, and the statistical significance of the association.

Table 2: Associations of birth weight with mother's age and weight at last menstrual period

	Correlation coefficient (r)	95% CI	P-value
Mother's Age	0.0899	-0.054, 0.23	0.2188
Weight at last menstrual period	0.1858	0.044, 0.32	0.0105

3. Univariate associations (200-400 words)

Examine the results of the regression models for birth weight presented in Table 3 below. The table displays the variable names and the results for the univariable and multivariable regression models, including β coefficients, 95% confidence intervals (CI) and p-values.

The variables included in the models are:

Outcome:

- Birth weight (*BWT*)

Exposure:

- History of hypertension coded as a binary factor variable (*ht_f*) Note: "No" is the reference category

Confounders:

- Mother's age in years (*AGE*)
- Smoking status during pregnancy coded as a binary factor variable (*smoke_f*)
- Mother's weight in pounds at last menstrual period (*LWT*)

Note: a separate model was fitted for each variable for the univariable results, and a single model containing all four variables was fitted for the multivariable results.

Please refer to the descriptions at the start for assignment for information about each variable to ensure your interpretation is appropriate for the scale of measurement.

- Interpret the beta coefficients and 95% confidence intervals from the univariable models for history of hypertension (*ht_f*), smoking status (*smoke_f*), mother's age (*AGE*) and mother's weight in pounds at last menstrual period (*LWT*) .
- For each univariable model, comment on the direction and magnitude of the estimated effect, and include an interpretation of the p-value.

4. Multivariable (adjusted) associations (100-200 words)

- Interpret the beta coefficients, 95% CIs, and p-values for each variable in the multivariable model in Table 3.
- Compare the coefficients with those in the univariable models and briefly describe what you observe.
- Additionally, comment on the magnitude of the effect of the primary exposure, hypertension, in terms of clinical relevance/importance.

Table 3: Linear regression estimates for low birth weight

	Univariable			Multivariable		
	Beta	95% CI	P-value	Beta	95% CI	P-value
Hypertension	-435.6	-858.3, -12.8	0.045	-579.5	-999.7, -159.2	0.008
Smoking	-281.7	-491.4, -72.1	0.009	-260.9	-464.9, -57.0	0.013
Age	12.4	-7.28, 32.0	0.219	5.50	-13.7, 24.7	0.574
Mother's weight	4.43	1.07, 7.79	0.010	5.17	1.75, 8.59	0.003

5. Diagnostics (100-200 words)

Examine Fig 5 and Fig 6 which show regression diagnostics for the multivariable model.

- Comment on the model fit.
- Are there sufficient diagnostics to judge the model fit? Why or why not?

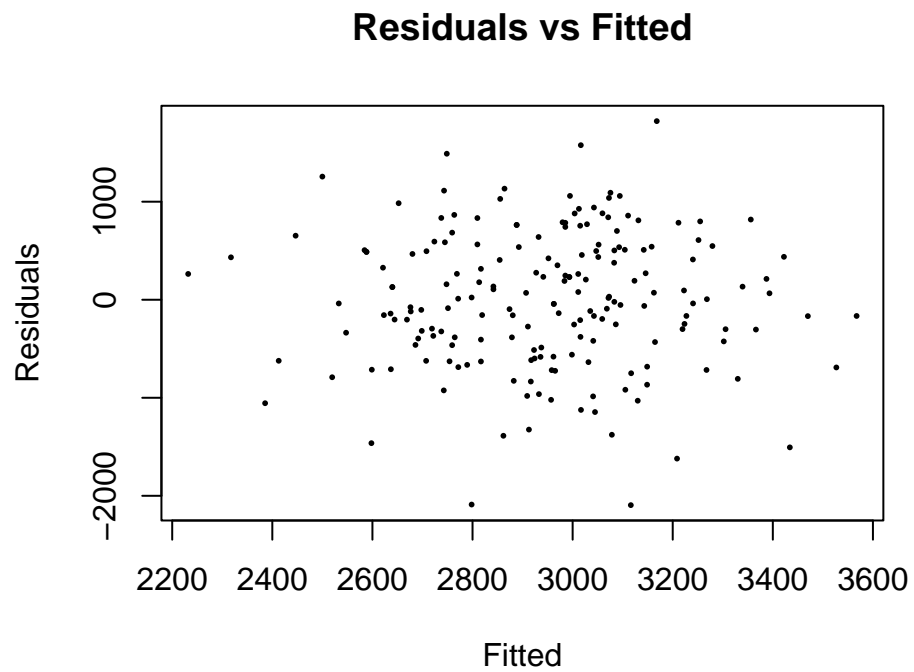


Figure 5: Residual vs fitted plot for the multivariable model

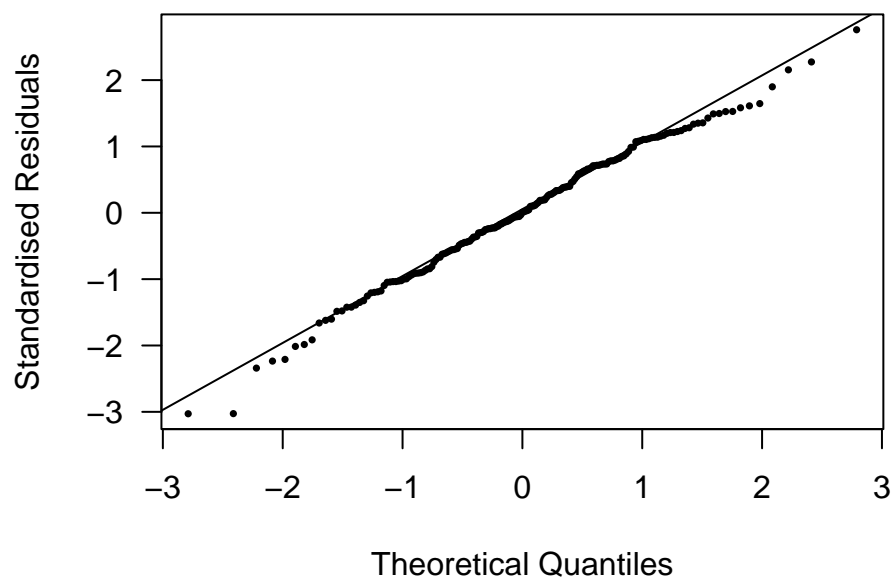


Figure 6: Q-Q plot for the multivariable model