



# LOGISTIC REGRESSION ANALYSIS OF ONSET PUBERTY GROWTH SPURT DATA

Christopher Morales <sup>†</sup> and Jayoung Kim<sup>†</sup>

Advisor : Jessica Jaynes, PhD<sup>‡</sup>

<sup>†</sup>Los Angeles City College

<sup>‡</sup>Department of Mathematics, California State University, Fullerton



## Abstract

Height and body mass index (BMI) have been shown to be correlated to earlier rates of childhood pubescence. Children who experience precocious, or early puberty, are seemingly prone to develop diseases such as breast cancer, ovarian cancer, and cardiovascular disease. In a recent research study conducted on Polish children [2], a multiple linear regression model was used to determine that height and BMI have an effect on the age of pubescence. In an effort to determine whether Polish children experienced on-time or early puberty, we propose the use of a logistic regression model by creating a binary categorization for the age of pubescence.

## Introduction

The purpose of this study is to examine the relationship between height and BMI with childhood pubescence. The data is from a previous study, which sampled 335 Polish children between the ages of 8 and 17 [2] from Wroclaw elementary schools. Of those 335 children, there are 162 boys and 173 girls. Variables considered in this study are shown in Table 1.

Onset age of Pubertal growth spurt (PGS)	Computed using Karlberg's infancy childhood puberty model
BMI SDS	Body mass index standard deviation score
Height gap SDS	Pre-pubertal height-SDS - Mid-parental height

Table 1: Variables related to childhood pubescence.

For our research, we wanted to create a binary categorization of the variable, PGS. The Polish children were divided into two groups, early puberty or on-time puberty, based on their age. According to U.S National Library of Medicine, the normal age range for female pubescence is 10-14 [1], while for males the normal age range is 12-16 [1]. Using these cut-points, if a polish child enters puberty before the normal range of pubescence, then it will be considered early puberty. To evaluate this, the predicted log-odds of early puberty will be calculated as a function of two variables: Height Gap, the difference between target height and the child's height at the onset of pubertal growth spurt (PGS), and BMI SDS.

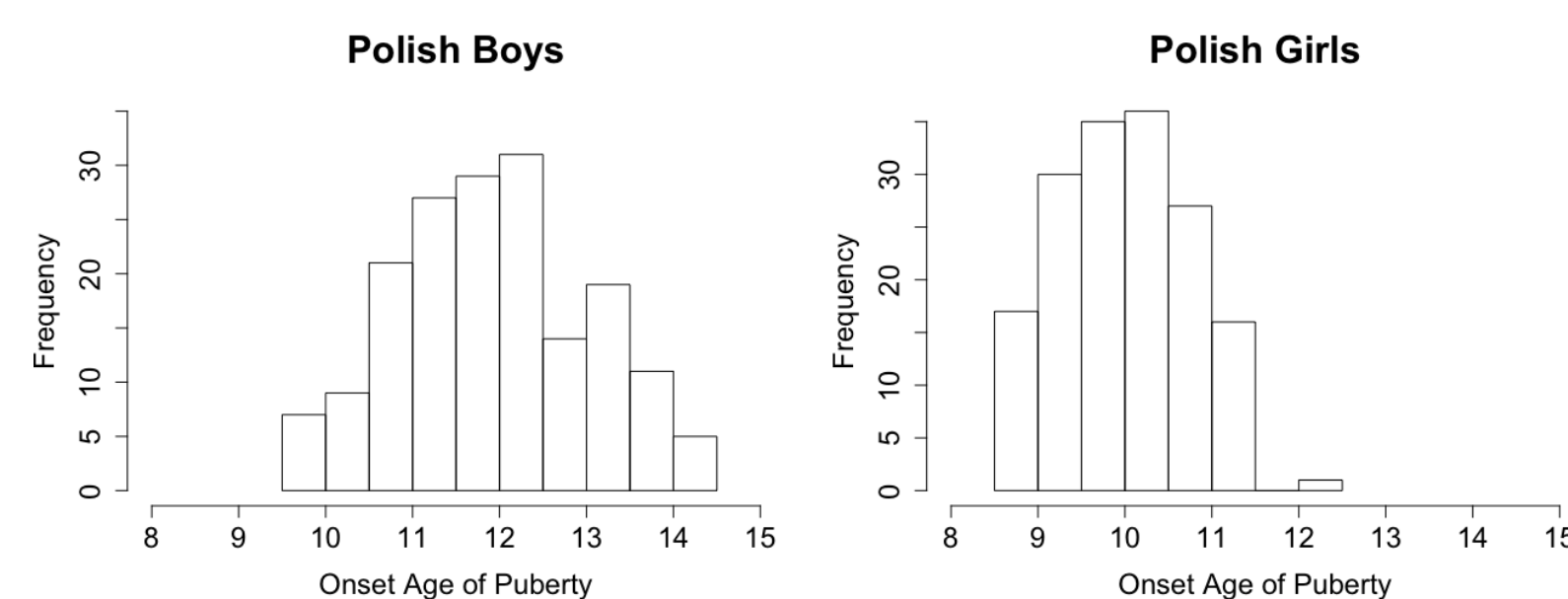


Fig. 1: Histogram of age of onset puberty

Figure 1 shows histograms of the average age of pubescence for the Polish children.

## Method

### Binary Categorization:

For  $i = 1, \dots, n$ , let  $y_i$  be the response of each child indicating on-time or precocious puberty:

$$y_i = \begin{cases} 0, & \text{on-time puberty} \\ 1, & \text{early puberty} \end{cases} \quad (1)$$

### Logistic Regression:

We model log-odds as the probability that a child reaches early puberty,

$$\log \left( \frac{P_i}{1 - P_i} \right) = \beta_{i0} + \beta_{i1}x_1 + \beta_{i2}x_2 + \dots + \beta_{ip}x_p,$$

- $P_i$  is the probability of early puberty
- The  $x_i$ 's in the logistic regression model are the predictor variables for  $y_i$ .
- The coefficients,  $\beta_i$ , of the logistic regression model are the log-odds of precocious puberty occurring.

## Model Fit

### Fitted Logistic Regression for Polish Boys:

$$\log \left( \frac{P_i}{1 - P_i} \right) = -0.12 + 0.61(\text{height.gap.SDS}) + 0.31(\text{BMI.SDS})$$

- The log odds of onset age of puberty increases by 0.61 for every one unit change in height gap SDS.
- The log odds of onset age of puberty increases by 0.31 for every one unit change in height gap SDS.

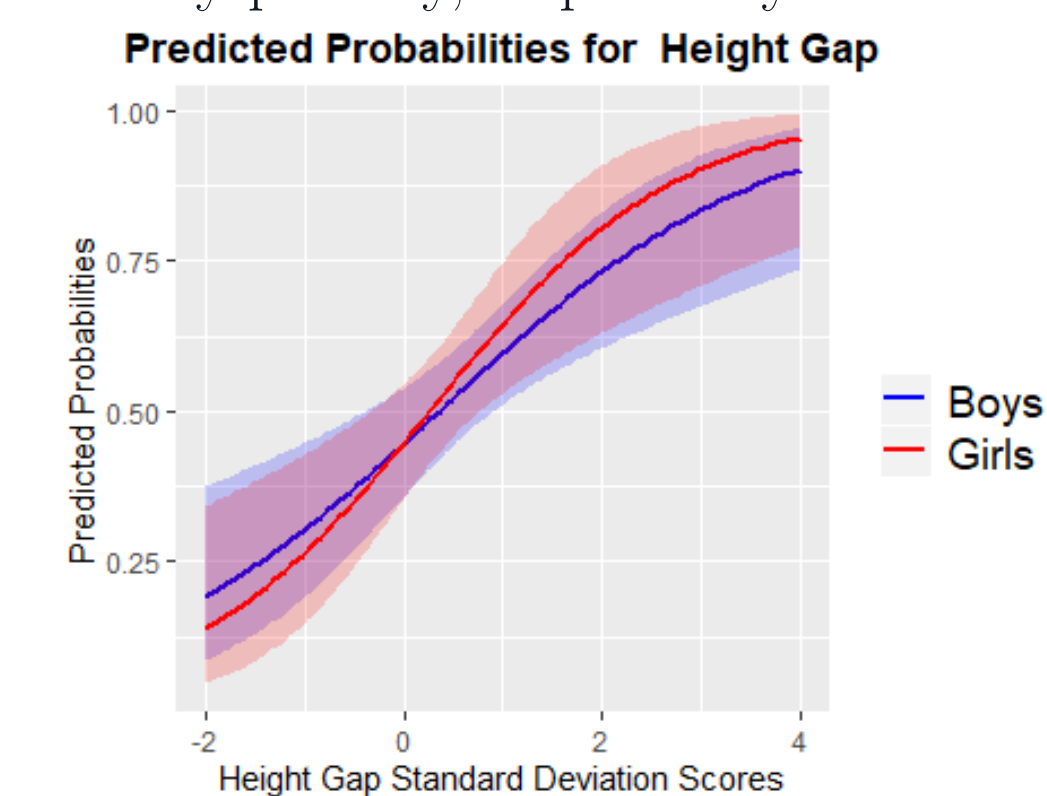
### Fitted Logistic Regression for Polish Girls:

$$\log \left( \frac{P_i}{1 - P_i} \right) = 0.43 + 0.81(\text{height.gap.SDS}) + 1.10(\text{BMI.SDS})$$

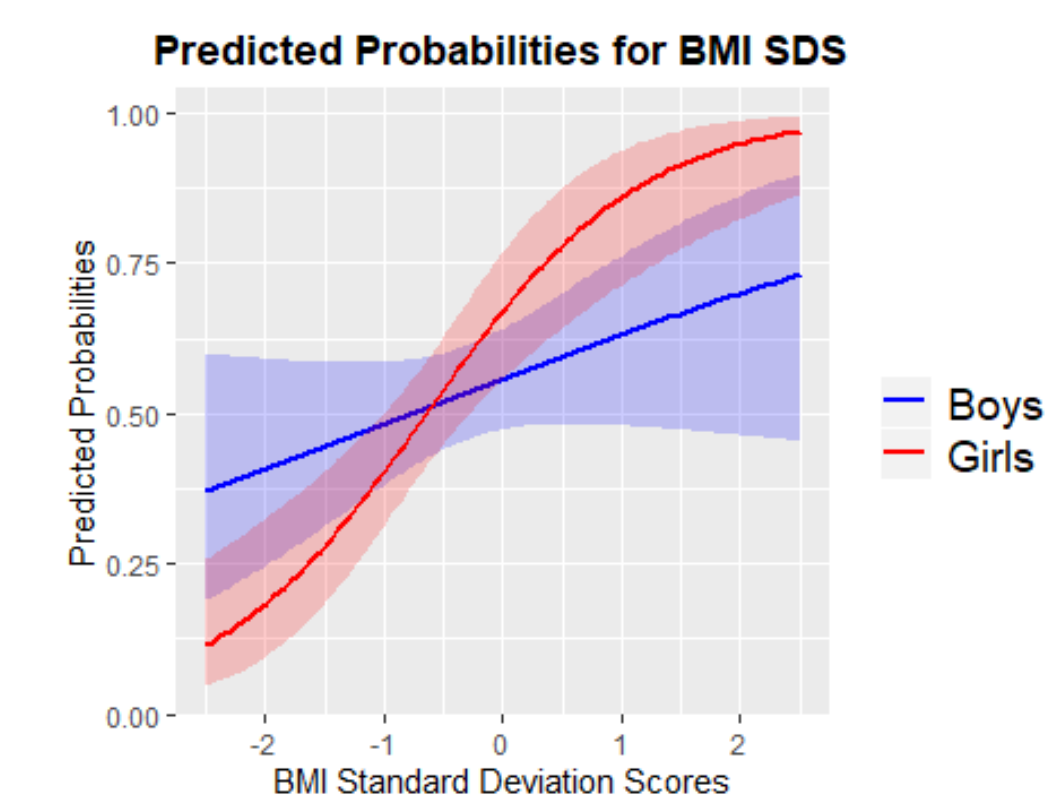
- The log odds of onset age of puberty increases by 0.81 for every one unit change in BMI SDS.
- The log odds of onset age of puberty] increases by 1.10 for every one unit change in BMI SDS.

## Predicted Probability Plots

Figures 2 and 3 compare the impact of height gap SDS and BMI SDS by gender. The lines and shaded regions indicate the estimated mean and 95% confidence interval of predicted probabilities of early puberty, respectively.



- Fig 2 shows the relationship between the predicted probabilities of early puberty and the values of height gap SDS for both boys and girls.



- Fig 3 shows the relationship between the predicted probabilities of early puberty and the values of BMI SDS for both boys and girls.

## Discussion

- The values of height gap SDS had a direct effect on the predicted probabilities of early puberty for both boys and girls.
- The values of BMI SDS had a sufficient impact on the predicted probabilities of early puberty for girls.
- The values of BMI SDS did not have a sufficient impact on the predicted probabilities of early puberty for boys.

## References and Acknowledgement

- [1] *Puberty*. Vol. 1. Medlineplus, U.S. National Library of Medicine, June 2019, p. 1.
- [2] S. Koziel Y. Limony and M. Friger. *Association between onset age of puberty and parental height*. Vol. 1. plus one, Jan. 2019, p. 1.