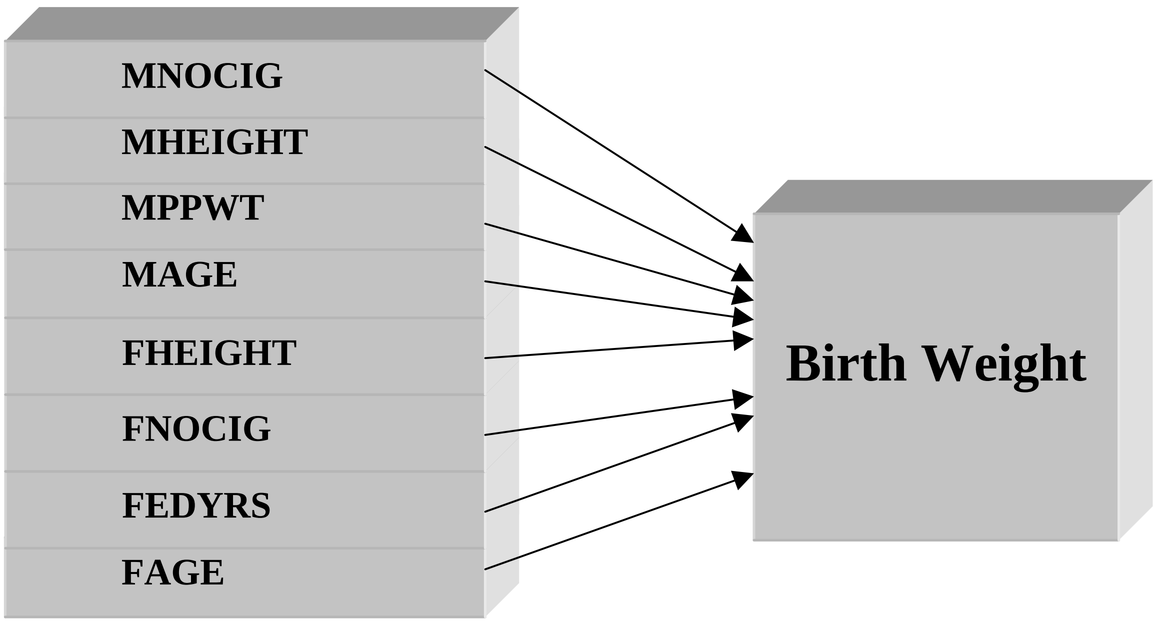
CHILD HEALTH AND DEVELOPMENT STUDY

10. Comparing the Influence of Maternal and Paternal Variables.

In this section we will use the child health and development data to evaluate the relative influences of the paternal versus the maternal variables on infant birth weight. We will address the problem by contrasting two separate regression analyses based on reduced models with the full model employing all nine variables. The comparison involves four maternal variables MNOCIG, MHEIGHT, MPPWT, and MAGE, and four paternal variables FHEIGHT, FNOCIG, FEDYRS, and FAGE.



The tentative full model for the problem developed in Section 7 has the form:

$$BWT = \beta_0 + \beta_1 * GESTWKS + \beta_2 * MNOCIG + \beta_3 * MAGE + \beta_4 * MHEIGHT +$$
$$\beta_5 * MPPWT + \beta_6 * FNOCIG + \beta_7 * FAGE + \beta_8 * FHEIGHT + \beta_9 * FEDYRS + ERROR.$$

The SPSS output for the full model is displayed below:

Multiple R	.51180
R Square	.26194
Adjusted R Square	.25203
Standard Error	.94472

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	9	212.22479	23.58053
Residual	**670**	**597.97074**	.89249

F = 26.42095 Signif F = .0000

Thus, the residual sum of squares for the full model is SSRes (Full) = 597.97074 with 670 degrees of freedom.

First, we investigate the impact of the paternal variables. In order to see the influence of the paternal variables, we test a null hypothesis that the paternal variables FHEIGHT, FNOCIG, FEDYRS, and FAGE can be removed from the full model. In other words, we test the null hypothesis

$$H_0 : \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0.$$

This hypothesis means that the paternal variables are not significant and can be removed from the model. The reduced regression model has now the form:

$$BWT = \beta_0 + \beta_1 * GESTWKS + \beta_2 * MNOCIG + \beta_3 * MAGE + \beta_4 * MHEIGHT + \beta_5 * MPPWT$$
$$+ ERROR.$$

Let us compare the change in the residual sum of squares. The SPSS output for the reduced model yields the residual sum of squares SSRes (Reduced) = 605.07071 with 674 degrees of freedom.

MULTIPLE LINEAR REGRESSION

Multiple R	.50317
R Square	.25318
Adjusted R Square	.24764
Standard Error	.94749

Analysis of Variance

	DF	**Sum of Squares**	**Mean Square**
Regression	5	205.12481	41.02496
Residual	**674**	**605.07071**	.89773

F = 45.69850 Signif F = .0000

Now we assess the significance of this increase with an F statistic:

$$F = \frac{\{SS\,\mathrm{Re}\,s(\mathrm{Re}\,duced) - SS\,\mathrm{Re}\,s(Full)\}/\{DF(\mathrm{Re}\,duced) - DF(Full)\}}{SS\,\mathrm{Re}\,s(Full)/DF(Full)}.$$

The F statistic follows an F-distribution with {DF(Reduced)-DF(Full)} as the numerator degrees of freedom and DF(Full) as the denominator degrees of freedom. The value of F is

$$F = \frac{(605.07071 - 597.97074)/(674 - 670)}{597.97074/670} = 1.989.$$

Under the null hypothesis that the paternal variables are not significant (have only a random influence on birth weight), the value of F is 1.989 and F follows an F-distribution with 674-670=4 and 670 degrees of freedom. This value produces a significance probability (p-value) of 0.095. The value of 0.05 is generally, but somewhat arbitrarily, used as the point above which variables are considered nonsignificant contributors. Using the criterion, we conclude that paternal variables are nonsignificant.

Now we examine the impact of the maternal variables. In order to see the influence of the maternal variables, we test a null hypothesis that the maternal variables MNOCIG, MHEIGHT, MPPWT, and MAGE can be removed from the full model. In other words, we test the null hypothesis

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0.$$

This hypothesis means that the maternal variables are not significant and can be removed from the model. The reduced regression model has now the form:

$$BWT = \beta_0 + \beta_1 * GESTWKS + \beta_6 * FNOCIG + \beta_7 * FAGE + \beta_8 * FHEIGHT + \beta_9 * FEDYRS + ERROR.$$

Let us compare the change in the residual sum of squares. The SPSS output for the reduced model yields the residual sum of squares SSRes (Reduced) = 645.65203 with 674 degrees of freedom.

MULTIPLE LINEAR REGRESSION

Multiple R	.45066
R Square	.20309
Adjusted R Square	.19718
Standard Error	.97874

Analysis of Variance

	DF	**Sum of Squares**	**Mean Square**
Regression	5	164.54350	32.90870
Residual	**674**	**645.65203**	.95794

F = 34.35359 Signif F = .0000

Now we assess the significance of this increase with an F statistic:

$$F = \frac{(645.65203 - 597.97074)/(674 - 670)}{597.97074/670} = 13.356.$$

Under the null hypothesis that the maternal variables are not significant (have only a random influence on birth weight), the value of F is 13.356 and F follows an F-distribution with 674-670=4 and 670 degrees of freedom. This probability of F that high when the null hypothesis is true (the maternal variables have no impact on birth weight) is less than 0.001, from an F distribution with 4 and 670 degrees of freedom. Thus there is strong reason to reject the null hypothesis and to claim that the maternal variables have significant influence on infant birth weight.

We summarize the changes in the residual sum of squares in the following table:

MODEL	SSRes	DF	CHANGE IN SSRes
FULL	597.97074	670	
PATERNAL REMOVED	605.07071	674	605.07071-597.97074=7.09997
MATERNAL REMOVED	645.65203	674	645.65203-597.97074=47.68129

In order to evaluate the relative impact on birth weight from the maternal and paternal variables, we compare the changes in the residual sums of squares for the reduced and the full model. The decrease in the residual sum of squares incurred by deleting the paternal variables is 7.1 and the decrease for the maternal variables is 47.7. Thus the relative influences of the maternal versus the paternal variables on infant birth weight can be estimated by the ratio 47.68129/7.09997=6.7157. The 6.7-fold increase in difference shows the considerably stronger influence of the maternal variables on an infant's birth weight.