MULTIPLE LINEAR REGRESSION

EXAMPLE: BLOOD VISCOSITY AND PACKED CELL VOLUME

The following blood viscosity data studied earlier are a good example of where multiple regression could be used. Recall that the data blood viscosity in samples taken from 32 hospital patients. We wish to model viscosity (y) as a function three covariates

- Packed Cell Volume (PCV), x_1.
- Plasma Fibrinogen, x_2.
- Plasma Protein, x_3.

Unit	Viscosity y	PCV x_1	Plasma Fib. x_2	Plasma Pro. x_3
1	3.71	40.00	344	6.27
2	3.78	40.00	330	4.86
3	3.85	42.50	280	5.09
4	3.88	42.00	418	6.79
5	3.98	45.00	774	6.40
6	4.03	42.00	388	5.48
7	4.05	42.50	336	6.27
8	4.14	47.00	431	6.89
9	4.14	46.75	276	5.18
10	4.20	48.00	422	5.73
11	4.20	46.00	280	5.89
12	4.27	47.00	460	6.58
13	4.27	43.25	412	5.67
14	4.37	45.00	320	6.23
15	4.41	50.00	502	4.99
16	4.64	45.00	550	6.37
17	4.68	51.25	414	6.40
18	4.73	50.25	304	6.00
19	4.87	49.00	472	5.94
20	4.94	50.00	728	5.16
21	4.95	50.00	716	6.29
22	4.96	49.00	400	5.96
23	5.02	50.50	576	5.90
24	5.02	51.25	354	5.81
25	5.12	49.50	392	5.49
26	5.15	56.00	352	5.41
27	5.17	50.00	572	6.24
28	5.18	47.00	634	6.50
29	5.38	53.25	458	6.60
30	5.77	57.00	1070	4.82
31	5.90	54.00	488	5.70
32	5.90	54.00	488	5.70

We consider four analyses:

$$\begin{aligned}
\textbf{Multiple regression :} \quad y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \\
\textbf{Regression on } \boldsymbol{x_1}\textbf{ :} \quad y &= \beta_0 + \beta_1 x_1 + \epsilon \\
\textbf{Regression on } \boldsymbol{x_2}\textbf{ :} \quad y &= \beta_0 + \beta_2 x_2 + \epsilon \\
\textbf{Regression on } \boldsymbol{x_3}\textbf{ :} \quad y &= \beta_0 + \beta_3 x_3 + \epsilon
\end{aligned}$$

Multiple Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.885[a]	.784	.761	.30370

a. Predictors: (Constant), Plasma Protein (g/100ml), Plasma Fibrinogen (mg/100ml), Packed Cell Volume (%)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.368	3	3.123	33.856	.000[a]
	Residual	2.582	28	.092		
	Total	11.950	31			

a. Predictors: (Constant), Plasma Protein (g/100ml), Plasma Fibrinogen (mg/100ml), Packed Cell Volume (%)

b. Dependent Variable: Blood Viscosity (cP)

Multiple Regression: Parameter Estimates

Tests are of the hypotheses
H0 : beta equal to 0
Ha : beta not equal to zero

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-1.378	.897		-1.537	.136	-3.215	.458
	Packed Cell Volume (%)	.117	.014	.839	8.584	.000	.089	.145
	Plasma Fibrinogen (mg/100ml)	.000	.000	.111	1.147	.261	.000	.001
	Plasma Protein (g/100ml)	.040	.097	.037	.412	.683	-.159	.239

a. Dependent Variable: Blood Viscosity (cP)

Only the packed cell volume coefficient is significantly different from zero ($p < 0.001$)

The other covariates do not seem to be significantly different from zero.

Regression on Packed Cell Volume only

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.879[a]	.772	.765	.30116

a. Predictors: (Constant), Packed Cell Volume (%)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.230	1	9.230	101.764	.000[a]
	Residual	2.721	30	.091		
	Total	11.950	31			

a. Predictors: (Constant), Packed Cell Volume (%)

b. Dependent Variable: Blood Viscosity (cP)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-1.223	.584		-2.094	.045	-2.416	-.030
	Packed Cell Volume (%)	.122	.012	.879	10.088	.000	.098	.147

a. Dependent Variable: Blood Viscosity (cP)

PCV is a significant term in the model ($p < 0.001$)

Regression on Plasma Protein only

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.457[a]	.209	.183	.56129

a. Predictors: (Constant), Plasma Fibrinogen (mg/100ml)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.499	1	2.499	7.932	.009[a]
	Residual	9.451	30	.315		
	Total	11.950	31			

a. Predictors: (Constant), Plasma Fibrinogen (mg/100ml)

b. Dependent Variable: Blood Viscosity (cP)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	3.871	.292		13.236	.000	3.274	4.468
	Plasma Fibrinogen (mg/100ml)	.002	.001	.457	2.816	.009	.000	.003

a. Dependent Variable: Blood Viscosity (cP)

Plasfib is a significant term in the model (p = 0.009)

Regression on Plasma Fibrinogen only

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.101[a]	.010	-.023	.62791

a. Predictors: (Constant), Plasma Protein (g/100ml)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.122	1	.122	.310	.582[a]
	Residual	11.828	30	.394		
	Total	11.950	31			

a. Predictors: (Constant), Plasma Protein (g/100ml)

b. Dependent Variable: Blood Viscosity (cP)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	5.296	1.174		4.510	.000	2.898	7.694
	Plasma Protein (g/100ml)	-.110	.198	-.101	-.556	.582	-.515	.295

a. Dependent Variable: Blood Viscosity (cP)

Plaspro is not a significant term in the model (p =0.582)