Regression Analysis Student Project
Winter 2016 Session

The Effect of Age and Weight on a Person's Systolic Blood Pressure

Introduction

Many cases of deaths worldwide are caused by high blood pressure, described by the National Institutes of Health as a common disease in which blood flows through blood vessels at higher than normal pressures. While it is known that this disease is more common to older people, there are also other factors causing the disease that are being considered.

This project intends to examine how one's age and weight affect his or her systolic blood pressure by creating a regression equation relating the two factors (age and weight) to the variable of concern, the systolic blood pressure.

Data

This study uses the available data from the Houghton Mifflin website found at the following link:

http://college.cengage.com/mathematics/brase/understandable_statistics/7e/students/datasets/mlr/frames/frame.html

Systolic Blood Pressure	Age (in years)	Weight (in pounds)
132	52	173
143	59	184
153	67	194
162	73	211
154	64	196
168	74	220
137	54	188
149	61	188
159	65	207
128	46	167
166	72	217

The dependent variable in the study is the systolic blood pressure which will be denoted by Y in the succeeding parts of this paper. It is the top number in a blood pressure reading which represents the amount of blood being exerted in the arteries when the heart is beating.

The two explanatory variables for which the dependent variable will be studied on are the person's age (measured in years) and weight (measured in pounds). Age will be denoted by X_1 while weight will be denoted by X_2.

Model

Three different models are done in the analyses. The first model is based on multiple regression method which analyzes the impact of both age and weight on the person's systolic blood pressure. The second and third models, on the other hand, are done based on the simple linear regression method and analyze the impact of age and weight, independently of each other, on the person's systolic blood pressure.

Microsoft Excel's Regression Analysis tool was utilized in coming up with the necessary regression values.

Model #1

The first model uses both the age and weight as explanatory variables for the person's systolic blood pressure.

The results of the regression analysis from Excel are as follows:

Table 1

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.988355758
R Square	0.976847104
Adjusted R Square	0.97105888
Standard Error	2.318211131
Observations	11

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1813.916268	906.9581341	168.7645659	0.000000287
Residual	8	42.99282278	5.374102848		
Total	10	1856.909091			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	30.99410295	11.94378039	2.594999401	0.031864509	3.451695983	58.53650992	3.451695983	58.53650992
Age (in years)	0.861414686	0.248231411	3.470208241	0.008440159	0.288992027	1.433837346	0.288992027	1.433837346
Weight (in pounds)	0.334859197	0.130668274	2.56266641	0.033507671	0.033537617	0.636180778	0.033537617	0.636180778

Looking at the results, it can be noted that both explanatory variables have positive slope. This means an increase in each variable results to a corresponding increase in systolic blood pressure.

While both variables, the age and the weight, have relatively low P-value, it can be noted that the P-value of the variable "age" is lower that of the variable "weight". This means that the impact of age on the person's diastolic blood pressure is more significant than the impact of weight.

The resulting regression equation based on Model #1 is:

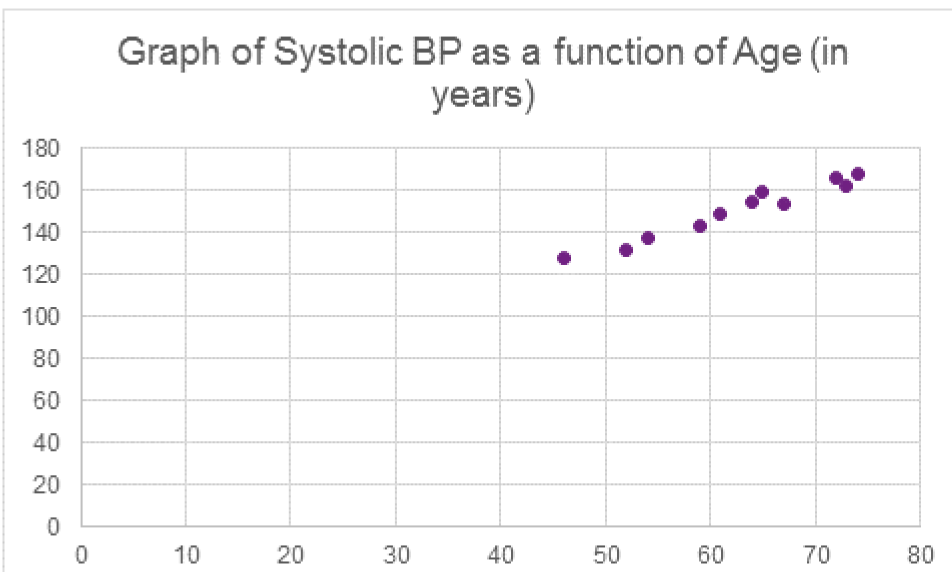
$$\mathbf{Y = 30.9941 + 0.8614X_1 + 0.3349X_2}$$

The above equation implies that for each year increase in a person's age, his systolic blood pressure increases by 0.8614, holding the weight constant. Another way of interpreting it is that for each pound increase in a person's weight, at the same age, his systolic blood pressure increases by 0.3349.

Model #2

The second model regresses the systolic blood pressure based on a person's age only.

Graph 1. Systolic BP as a function of Age (in years)



The results of the regression analysis from Excel are as follows:

Table 2

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.978693374
R Square	0.957840721
Adjusted R Square	0.953156356
Standard Error	2.949311573
Observations	11

ANOVA

	df	SS	MS	F	Significance F
Regression	1	1778.623142	1778.623142	204.4761356	0.00000017078
Residual	9	78.28594879	8.698438754		
Total	10	1856.909091			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	58.70551543	6.452374608	9.09828071	0.00000781	44.10922999	73.30180087	44.10922999	73.30180087
Age (in years)	1.463230466	0.102327278	14.29951522	0.00000017	1.231750082	1.694710851	1.231750082	1.694710851

Based on Graph 1, even without looking at the results of the regression analysis in Table 2, it is clear that there is a positive correlation between the age of a person and his systolic blood pressure. This means that as age increases (or simply, as a person becomes older), the systolic blood pressure tends to increase, too. Also, as the correlation coefficient's value is close to 1, it can be implied that there is a strong correlation between the two variables.

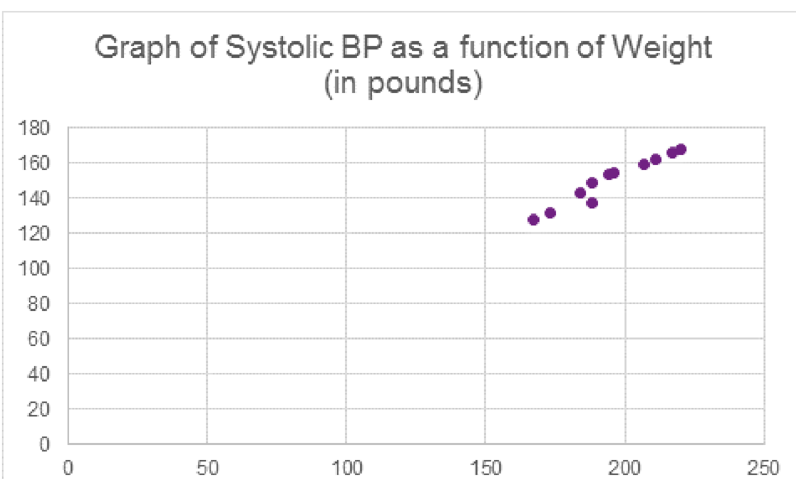
The resulting regression equation of Model #2 is:

$$\mathbf{Y = 58.7055 + 1.4632X_1}$$

Model #3

The third model regresses the systolic blood pressure on a person's weight.

Graph 2. Systolic BP as a function of Weight (in pounds)



The results of the regression analysis from Excel are as follows:

Table 3

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.970564376
R Square	0.941995209
Adjusted R Square	0.935550232
Standard Error	3.459441583
Observations	11

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1749.199466	1749.199466	146.1595958	0.0000007227
Residual	9	107.7096246	11.96773607		
Total	10	1856.909091			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1.141609558	12.36447433	0.092329809	0.928458547	-26.82877461	29.11199373	-26.82877461	29.11199373
Weight (in pounds)	0.763842562	0.063181537	12.08964829	0.00000072	0.620915995	0.906769128	0.620915995	0.906769128

Again, based on Graph 2, there is a positive correlation between the weight of a person and his systolic blood pressure. This means that as weight increases, the systolic blood pressure tends to increase, too. However, it can be noted that the correlation coefficient of Model #3 is lower than that of Model #2. This means higher correlation between the age and systolic blood pressure than between the weight and systolic blood pressure. This supports our observation in Model #1 that the impact of age on the systolic blood pressure is more significant that the impact of weight.

The resulting regression equation of Model #3 is:

$$\mathbf{Y = 1.1416 + 0.7638X_1}$$

Conclusion

The results of our analyses based on the three models indicate that there is indeed a strong positive correlation between the age and the weight of a person and his systolic blood pressure. However, based on the comparison of the three models, it is also noted that an increase in age affects the systolic blood pressure more than an increase in the weight does.

References:

http://www.nhlbi.nih.gov/health/health-topics/topics/hbp/