Statistics 512: Solution to Homework#7

Solution Key to Problem 1. It is unclear here whether you are to use the centered or uncentered variables. Here I only present the results using centered variables. When uncentered variables are used, many of the results are the same but some (the influence summaries) are slightly different.

First, I give the model output for the transformed model. The four residual plots are then shown. There does not appear to be any further changes to the model based on these except to notice that the variability does seem to get smaller as x2 increases. This may be due to some influential points.

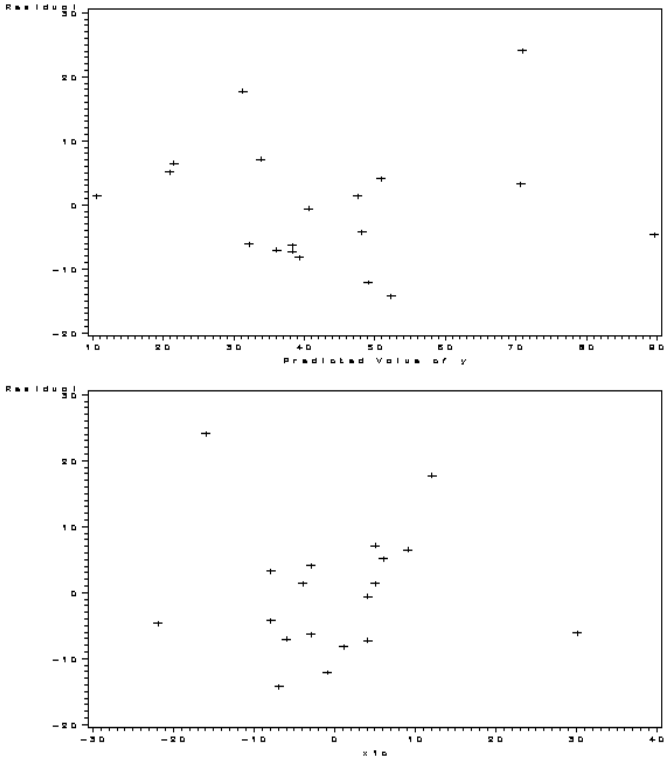
Analysis of Variance

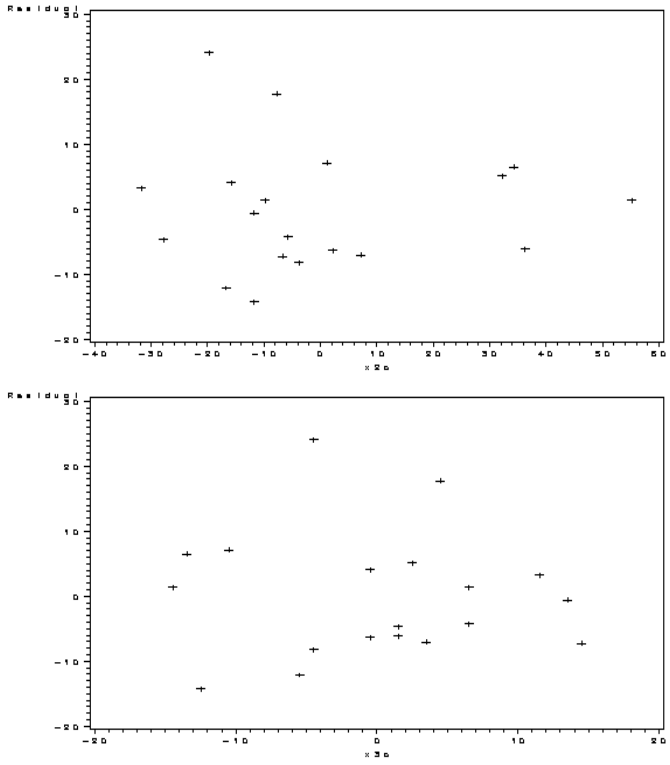
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	6407.21941	2135.73980	19.06	<.0001
Error	15	1680.46480	112.03099		
Corrected Total	18	8087.68421			

Root MSE	10.58447	R-Square	0.7922
Dependent Mean	43.26316	Adj R-Sq	0.7507
Coeff Var	24.46532		

Parameter Estimates

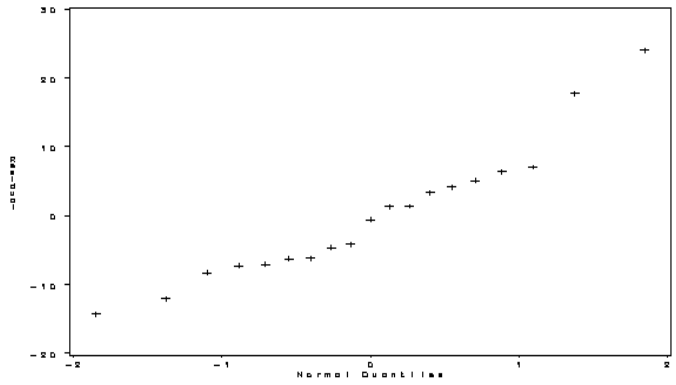
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	37.88818	2.85220	13.28	<.0001	0
x1c	1	-0.67453	0.29678	-2.27	0.0382	1.75458
x2c	1	-0.60239	0.14388	-4.19	0.0008	1.82193
x1cx2c	1	0.03335	0.00928	3.59	0.0027	1.13073





It is very clear through the VIF (above) that the centering takes care of multicollinearity.

The normality assumption appears reasonable. The normal qqplot is shown below. The correlation between the expected values and the ordered values is 0.96338. This means there is not enough evidence to reject the null.



The studentized deleted residuals are shown below. They are the same for both models. Since there are 19 observations, we'd reject if the abs(residual) were larger than 3.649. This is based on 15-1=14 degrees of freedom and the Bonferonni correction.

The hat values are also shown below. These are also the same in both models. The rule of thumb is to see if any value is greater than 2(4)/19=0.421. In this case, observation 3, 8, and 15 appear influential.

The DFbetas are also shown below. In looking whether there are values greater than 1, we see that observation 3 is ok, observation 7 has some impact on x1c, observation 8 has impact on x1c, x2c, and x1cx2c, and observation 15 is ok. In the untransformed case, observation 3 looks ok, observation 7

influences both the intercept and x1. Observation 8 influences all the parameters and observation 15 is ok.

The DFFITS and Cook's D is not shown but for observation 8, the values are -4.7798 and 4.991 respectively. These suggest concern.

			Hat Diag	Cov		DFBETAS			
Obs	Residual	RStudent	H	Ratio	DFFITS	Intercept	x1c	x2c	x1cx2c
1	17.7397	2.2095	0.2757	0.5499	1.3632	0.8195	1.0978	-0.6979	-0.5955
2	4.1605	0.3989	0.0834	1.3741	0.1203	0.0923	0.0248	-0.0594	-0.0209
3	-4.6164	-0.6292	0.5389	2.5561	-0.6802	0.0718	0.3822	0.0726	-0.4819
4	-6.2590	-0.6049	0.0848	1.2988	-0.1842	-0.1722	0.0558	-0.0706	0.0929
5	5.0963	0.5172	0.1757	1.4821	0.2387	0.1334	-0.0587	0.1857	-0.0422
6	6.4898	0.6618	0.1737	1.4100	0.3035	0.1367	-0.0320	0.2039	0.0104
7	24.0398	3.3141	0.2178	0.1661	1.7486	0.2873	-1.0884	-0.0756	0.8475
8	-6.1424	-1.7794	0.8783	4.7895	-4.7798	0.7275	-2.4073	1.0195	-3.2858
9	3.3135	0.3379	0.1925	1.5799	0.1650	0.0367	0.0148	-0.1168	0.0701
10	-12.0731	-1.2232	0.1017	0.9773	-0.4116	-0.3054	-0.1630	0.2391	0.1017
11	-7.2550	-0.7153	0.1116	1.2849	-0.2534	-0.2103	-0.1407	0.0974	0.1183
12	1.3548	0.1282	0.0680	1.4073	0.0346	0.0309	-0.0024	-0.0056	-0.0095
13	-14.3075	-1.4574	0.0753	0.8100	-0.4159	-0.3130	0.1345	0.0381	0.0323
14	7.1013	0.6921	0.0929	1.2699	0.2215	0.1995	0.1044	-0.0314	-0.1097
15	1.4369	0.1821	0.4798	2.5095	0.1749	0.0575	-0.0835	0.1617	-0.0157
16	-4.1795	-0.4021	0.0897	1.3826	-0.1262	-0.0972	0.0694	-0.0338	0.0283
17	-7.0614	-0.7092	0.1444	1.3375	-0.2914	-0.2245	0.1624	-0.1828	0.1426
18	-0.5824	-0.0573	0.1391	1.5292	-0.0230	-0.0172	-0.0141	0.0119	0.0099
19	-8.2559	-0.8021	0.0768	1.1926	-0.2314	-0.2230	-0.0620	0.0332	0.1143

Sum of Residuals	0
Sum of Squared Residuals	1680.46480
Predicted Residual SS (PRESS)	5102.49435

Output Statistics

			Hat Diag	Cov		DFBETAS			
Obs	Residual	RStudent	H	Ratio	DFFITS	Intercept	x1	x2	x1x2
1	17.7397	2.2095	0.2757	0.5499	1.3632	-0.7472	1.0870	0.2239	-0.5955
2	4.1605	0.3989	0.0834	1.3741	0.1203	0.0072	0.0304	-0.0059	-0.0209
3	-4.6164	-0.6292	0.5389	2.5561	-0.6802	-0.6519	0.5919	0.4334	-0.4819
4	-6.2590	-0.6049	0.0848	1.2988	-0.1842	0.0406	-0.0405	-0.1059	0.0929
5	5.0963	0.5172	0.1757	1.4821	0.2387	-0.0487	-0.0006	0.1089	-0.0422
6	6.4898	0.6618	0.1737	1.4100	0.3035	-0.0276	-0.0263	0.0719	0.0104
7	24.0398	3.3141	0.2178	0.1661	1.7486	1.4541	-1.2776	-0.7415	0.8475
8	-6.1424	-1.7794	0.8783	4.7895	-4.7798	-1.5469	1.1866	3.1623	-3.2858
9	3.3135	0.3379	0.1925	1.5799	0.1650	0.1021	-0.0461	-0.1051	0.0701
10	-12.0731	-1.2232	0.1017	0.9773	-0.4116	0.0359	-0.1717	0.0092	0.1017
11	-7.2550	-0.7153	0.1116	1.2849	-0.2534	0.1089	-0.1720	-0.0608	0.1183
12	1.3548	0.1282	0.0680	1.4073	0.0346	0.0013	0.0060	0.0058	-0.0095
13	-14.3075	-1.4574	0.0753	0.8100	-0.4159	-0.1260	0.0513	-0.0120	0.0323
14	7.1013	0.6921	0.0929	1.2699	0.2215	-0.1075	0.1446	0.0797	-0.1097
15	1.4369	0.1821	0.4798	2.5095	0.1749	-0.0155	-0.0353	0.0771	-0.0157
16	-4.1795	-0.4021	0.0897	1.3826	-0.1262	-0.0227	0.0174	-0.0372	0.0283
17	-7.0614	-0.7092	0.1444	1.3375	-0.2914	0.0519	-0.0186	-0.1920	0.1426
18	-0.5824	-0.0573	0.1391	1.5292	-0.0230	0.0091	-0.0157	-0.0036	0.0099
19	-8.2559	-0.8021	0.0768	1.1926	-0.2314	0.0806	-0.1241	-0.0828	0.1143

Sum of Residuals	0
Sum of Squared Residuals	1680.46480
Predicted Residual SS (PRESS)	5102.49435

Solution Key to Problem 2. Omitted.