Psychology 407

Assignment C

In an experiment designed to assess a possible curvilinear relationship between level of background noise on task performance, an experimenter assigned (at random) 10 subjects to each of 6 noise levels (assumed to be equally-spaced values of 1, 2, ..., 6) and obtained a "number correct" score on a heavily speeded performance measure. The data for this study turned out as follows (the columns are labeled by Noise Level):

one	two	three	four	five	six
18	34	39	37	15	14
24	36	41	32	18	19
20	39	35	25	27	5
26	43	48	28	28	25
23	48	44	29	22	7
29	28	38	31	24	13
27	30	42	34	21	10
33	33	47	38	19	16
32	37	53	43	13	20
38	42	33	23	33	11

Because of possible computational issue in fitting polynomial models (question: what are they?), the 6 noise levels will be coded as deviations from the mean noise level (a value of 3.5); thus, the 6 noise levels are actually -2.5, -1.5, -.5, +.5, +1.5, +2.5. These latter deviation values should be assumed in *everything* that follows.

Summary Information on Performance (the standard deviation is based on an unbiased variance estimate):

Noise level	Sample Size	Mean	Standard Deviation
-2.5	10	27.0	6.164
-1.5	10	37.0	6.164
-.5	10	42.0	6.164
+.5	10	32.0	6.164
+1.5	10	22.0	6.164
+2.5	10	14.0	6.164
Overall	60	29.0	11.087

(*Any* indication that these data are "made up"?)

The end two pages give SYSTAT results on fitting a variety of polynomial models. Here, PER stands for performance and NOISED stands for noise deviated from the mean.

Questions:

a) Plot the data: noise against performance. Indicate on the plot the mean performance level within each noise level.

b) Replot just the mean performance levels within each noise level and on this graph represent all *five* linear/curvilenear functions given by the SYSTAT output. Comment on what appears to provide a "reasonable" fit.

c) Calculate a "pure error" sum-of-squares from the summary information provided for performance. What would a plot in (b) look like if a polynomial of order 5 were fitted? And what would be the residual sum-of-squares? Provide the analysis-of-variance table for fitting the order 5 polynomial. (If in a previous life you studied one-way analysis-of-variance, comment on the correspondence between the last table you gave and what would be usually provided in the one-way analysis of variance context.)

d) Obtain the "extra" sums-of-squares indicated (here, X is the noise level):

SSR(X); SSR(X^2 | X);
SSR(X^3 | X, X^2);
SSR(X^4 | X, X^2, X^3);

$SSR(X^5 \mid X, X^2, X^3, X^4)$; and
$SSR(X^3, X^4, X^5 \mid X, X^2)$
$SSR(X^4, X^5 \mid X, X^2, X^3)$.

Test whether there is a significant lack-of-fit for a second order and for a third order model using the "pure error" term — give the two corresponding analysis-of-variance tables. Comment on how these tests relate to the intuition you provided in (b).

e) What is the relation between all of the residual mean squares generated in the SYSTAT analyses and the mean square for pure error? Are they all estimates of error? In what sense and under what conditions?

f) Look at the SYSTAT analysis for the third order model. Show numerically how the test for the coefficient on X^3 can be generated using the extra sum of squares principle. In carrying out this test, what assumption is being made about the residual mean-squares for the third order model.

g) Look at the SYSTAT analysis for the second order model. What do the given tolerances tell you about the relation between X and X^2? Why should this relation hold here for our data? Comment on the change or lack of change in the regression coefficients as the order of the model increases. How general would you expect such a result to be when other data sets are considered?

Dep Var: PERFORMANCE N: 60 Multiple R: 0.0 Squared multiple R: 0.0

Adjusted squared multiple R: 0.000 Standard error of estimate: 11.087

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	29.000	1.431	0.0	.	20.261	0.000

Dep Var: PERFORMANCE N: 60 Multiple R: 0.533 Squared multiple R: 0.284

Adjusted squared multiple R: 0.272 Standard error of estimate: 9.461

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	29.003	1.221	0.0	.	23.745	0.000
NOISED	-3.430	0.715	-0.533	1.000	-4.797	0.000

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	2060.151	1	2060.151	23.015	0.000
Residual	5191.849	58	89.515		

Dep Var: PERFORMANCE N: 60 Multiple R: 0.809 Squared multiple R: 0.655

Adjusted squared multiple R: 0.643 Standard error of estimate: 6.628

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	36.837	1.317	0.0	.	27.962	0.000
NOISED	-3.426	0.501	-0.533	1.000	-6.842	0.000
NOISED *NOISED	-2.684	0.343	-0.609	1.000	-7.821	0.000

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	4747.738	2	2373.869	54.032	0.000
Residual	2504.262	57	43.934		

Dep Var: PERFORMANCE N: 60 Multiple R: 0.838 Squared multiple R: 0.703

Adjusted squared multiple R: 0.687 Standard error of estimate: 6.201

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	36.832	1.232	0.0	.	29.885	0.000
NOISED	-7.145	1.317	-1.110	0.127	-5.425	0.000
NOISED *NOISED	-2.682	0.321	-0.608	1.000	-8.356	0.000
NOISED *NOISED *NOISED	0.736	0.244	0.618	0.127	3.021	0.004

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	5098.694	3	1699.565	44.200	0.000
Residual	2153.306	56	38.452		

Dep Var: PERFORMANCE N: 60 Multiple R: 0.843 Squared multiple R: 0.710

Adjusted squared multiple R: 0.689 Standard error of estimate: 6.182

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	38.091	1.642	0.0	.	23.199	0.000
NOISED	-7.142	1.313	-1.110	0.127	-5.439	0.000
NOISED *NOISED	-4.369	1.494	-0.991	0.046	-2.925	0.005
NOISED *NOISED *NOISED	0.736	0.243	0.618	0.127	3.029	0.004
NOISED *NOISED *NOISED *NOISED	0.249	0.215	0.392	0.046	1.156	0.253

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	5149.747	4	1287.437	33.682	0.000
Residual	2102.253	55	38.223		
