August 24, 2009

Statistics 621 Waiver Exam

Read these instructions carefully.

Before the exam begins, on the answer sheet that you were given...

- **Use a #2 pencil only**. **Do *not* use a pen.**
  Erase any changes on the form completely. Do not damage or fold the page.
- **Write in your name and Penn student id number**.
  Your Penn ID number appears in bold numerals on your ID card.
- **Fill in the "bubbles" under the letters of your name *and* student id number.**
  If you don't, the computer will not record your name or ID number.

After the exam begins …

- **Choose one *best* answer for each question.**
  Picking more than one answer is an error.
- **You may refer to 1 page of handwritten notes during the exam.**
  No other references are permitted. You can use the front and back of the page.
- **You may use a calculator. Phones or other electronic devices are not allowed.**
  If you have a phone with you, please turn it off prior to the exam.

You have **two hours** for the exam. The JMP output associated with one or more questions should be considered an essential part of the question. The word "significant" means "statistically significant". SRM stands for Simple Regression Model; MRM stands for multiple regression model.

Turn in only the solution page. Keep the exam itself so that you may compare your answers and mark your choices on your copy of the exam. (Regardless of what you write on your copy of the exam, however, only the answers marked on the answer form will be considered.)

Your score is the number of correct answers. The questions are equally weighted. Some questions may be dropped and not counted as part of the overall score. There is no deduction for incorrect answers. Solutions will be posted in Web Café. Use the "My Grades" feature of Web Café to find your score.

STOP

*Do **not** turn the page until you are instructed to proceed.*

1. Mark the answer to Question 1 on your answer form "a".
   (This question identifies your version of the exam. Mark the answer form as indicated.)

2. The standard error of the estimated slope in the summary of a simple least squares regression equation
   a) Estimates the sample-to-sample standard deviation of the estimated slope.
   b) Increases with the size of the r^2 statistic of the fitted regression.
   c) Decreases as the variability of the explanatory variable decreases.
   d) Is unreliable unless the residuals are exactly normally distributed.
   e) Increases with the size of the sample.

3. The RMSE of a fitted simple regression model
   a) Decreases when an explanatory variable is added to the regression model.
   b) Estimates the standard deviation of the explanatory variable.
   c) Is ¼ of the width of the approximate 95% confidence interval for the slope.
   d) Is ½ of the width of the approximate 95% prediction interval.
   e) Estimates the standard deviation of ε in the Simple Regression Model.

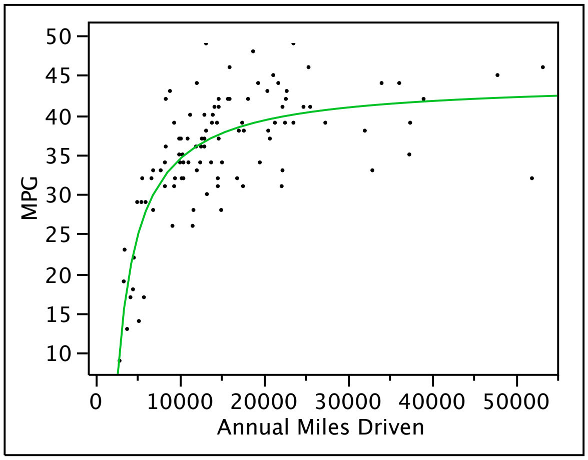
4. A key limitation of the R^2 statistic in regression analysis is that it
   a) Increases when the size of the sample grows.
   b) Increases when any explanatory variable is added to the regression.
   c) Is unrelated to the residual SD of the regression model.
   d) Is only interpretable as the square of a correlation in simple regression.
   e) Does not reflect the influence of leveraged outliers.

5. The presence of heteroscedasticity in an estimated regression model
   a) Implies that 95% confidence intervals for the estimated slope are too narrow.
   b) Suggests that the estimated slope is larger than the population slope.
   c) Implies that the estimated intercept of the fitted model is unreliable.
   d) Is the result of a small subset of highly leveraged observations.
   e) Indicates that 95% prediction intervals from the model are unreliable.

6. The most useful plot for identifying the presence and size of autocorrelation in an estimated regression model is
   a) A time sequence plot of the response variable.
   b) A scatterplot of the residuals versus the explanatory variable.
   c) A time sequence plot of the explanatory variable.
   d) A scatterplot of the response versus the explanatory variable.
   e) A scatterplot of the residuals versus the lagged residuals.

7. If the residuals from a fitted least squares regression are not normally distributed, then
   a) The 95% confidence interval for the intercept is longer than it should be.
   b) The standard error of the slope underestimates the SD of the slope estimator.
   c) The response variable in the regression should be transformed to a log scale.
   d) The 95% prediction intervals may be too long or too short.
   e) The fitted model omits an important explanatory variable.

Questions 8-14. The "Cash for Clunkers" program (formally CARS, Car Allowance Rebate System) encouraged consumers to trade in older, inefficient cars for new, more efficient cars by contributing a cash payment toward the value of a traded-in vehicle. Participating consumers filled in a survey at the time of purchase. This question considers data from n = 105 survey forms. *MPG* is the miles per gallon rating of the purchased car, and *Annual Miles Driven* is the number of miles driven annually by the consumer making the purchase.



Transformed Fit to Reciprocal

MPG = 44.187601 - 98356.094*Recip(Annual Miles Driven)

Summary of Fit

RSquare	0.615834
RSquare Adj	0.612104
Root Mean Square Error	4.852591
Mean of Response	35.28571
Observations (or Sum Wgts)	105

Lack Of Fit

Analysis of Variance

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	44.187601	0.839166	52.66	<.0001*
Recip(Annual Miles Driven)	-98356.09	7654.392	-12.85	<.0001*

8. The equation of the estimated model implies that
   a) The number of miles driven annually is not associated with MPG.
   b) Cars purchased during this program average about 44.2 MPG.
   c) People who drive farther than 50,000 miles purchase cars with lower MPG on average.
   d) Few customers purchase cars that on average obtain high MPG.
   e) Customers who drive more than 50,000 miles purchase cars averaging nearly 44.2 MPG.

9. The shown Root Mean Square Error (4.85), assuming the SRM, indicates that
   a) 95% of customers drive between about 25,000 to 45,000 miles annually.
   b) The standard deviation of *MPG* is 4.85.
   c) The variance of the underlying model errors is 4.85 *MPG*.
   d) Residuals of magnitude greater than 4.85 *MPG* are important outliers.
   e) 95% of prediction errors when using this equation are between ± 9.7 *MPG*.

10. Anne and Dave purchase cars in this program. Anne drives 10,000 miles farther than Dave annually. The difference in predicted miles per gallon of the car purchased by Anne compared to the car purchased by Dave (use the shown model and assume the SRM holds)
    a) Is about 9.8 ± 1.5 MPG. (Anne's car has higher MPG.)
    b) Is about 9.8 ± 1.5 MPG. (Dave's car has higher MPG.)
    c) Depends on how far Dave drives annually.
    d) Is about 9.8 ± 9.7 MPG. (Anne's car has higher MPG.)
    e) Is about 9.8 ± 9.7 MPG. (Dave's car has higher MPG.)

11. To obtain a more accurate estimate of the slope in the fitted equation, 100 more observations are to be added from the same population as these. The greatest improvement in accuracy will be obtained by adding observations that
    a) Lie close to the estimated equation of this regression.
    b) Measure characteristics of high-mileage drivers.
    c) Measure characteristics of drivers who bought cars that obtain high MPG.
    d) Add variation to the explanatory variable.
    e) Describe customers who have similar driving habits.

As an alternative to the model shown on the previous page, a revised equation was fit to the original sample of 105 observations. The estimated revised equation is

$$\textit{Estimated MPG} = -50.595 + 9.04 \log_e(\textit{Annual Miles Driven})$$

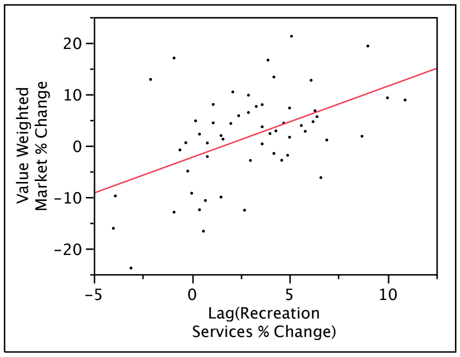
which has $R^2 = 0.402$.

12. The standard deviation of the residuals of the revised equation
    a) Is less than 4.85 MPG.
    b) 8.46 MPG
    c) 6.05 MPG
    d) 6.89 MPG
    e) 7.43 MPG

13. The best interpretation of the slope in the revised equation is
    a) The elasticity of *MPG* with respect to annual miles driven is 9.04.
    b) Average *MPG* increases about 0.09 per 1% increase in annual miles driven.
    c) On average, a 1% increase in annual miles driven increases *MPG* by 9.04%.
    d) On average, each additional mile driven increases *MPG* by about 0.09.
    e) Average *MPG* increases about 9% for each 1,000 mile increase in annual miles driven.

14. The revised equation
    a) Has a lower R^2 than the original equation, indicating that the *original* equation fits better.
    b) Has a lower R^2 than the original equation, indicating that the *revised* equation fits better.
    c) Cannot be compared to the original equation because it uses a different transformation.
    d) Explains about the same proportion of variation in *MPG* of purchased cars.
    e) Produces narrower 95% prediction intervals for high-mileage drivers.

Questions 15-26. A day trader who regularly invests in stock built a regression model to help with his investing. He regressed quarterly percentage change in the stock market (measured by the value-weighted index, grouped into 3 month periods) on the percentage change in personal consumption expenditures for recreational services. Recreational services include purchases of tickets to movies, amusement parks and spectator sports and well as purchases of books, magazines, and audio/video sources (such as music or DVDs). The quarterly data span 1995 through the first quarter of 2009.



Summary of Fit

RSquare	0.237284
RSquare Adj	0.22316
Root Mean Square Error	8.17164
Mean of Response	1.77082
Observations (or Sum Wgts)	56

Lack Of Fit

Analysis of Variance

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2.313449	1.478302	-1.56	0.1234
Lag(Recreation Services % Change)	1.3844978	0.337786	4.10	0.0001*

15. The correlation between percentage changes in the value weighted market index and percentage changes in the lag of percentage changes in purchases of recreational services is about
    a) 0.24
    b) 0.49
    c) 0.00
    d) 0.14
    e) Not determined by the shown output.

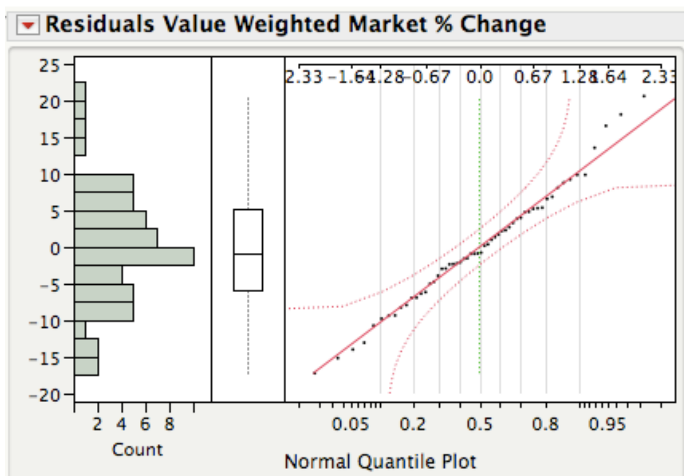
16. A theoretical model developed by the trader implies that expected percentage changes in the market in the current quarter are proportional to percentage changes in purchases of recreational services in the prior quarter. Assuming the SRM holds, the shown regression model
    a) Is consistent with this aspect of the model developed by the trader.
    b) Rejects this aspect of the model developed by the trader at the $\alpha = 0.05$ level.
    c) Rejects this aspect of the model developed by the trader, but only at the $\alpha = 0.1234$ level.
    d) Does not address this aspect of the model developed by the trader.
    e) Fits too poorly to reject or confirm the model developed by the trader.

17. Consider quarters in which purchases of recreational services do not change. The fitted model
    a) Implies that the market does not change during following quarters.
    b) Implies that the market falls about 2.3% during following quarters.
    c) Implies that the market grows about 1.8% during following quarters.
    d) Implies that the market grows about 1.4% during following quarters.
    e) Only describes changes in the market when purchases of recreational services change.

18. Purchases of recreational services fell 0.9 percent during the *first* quarter of 2009. The shown model, assuming the SRM holds, predicts the percentage change in the market during the *second* quarter of 2009
    a) Will be between -19.9% to 12.8%, with 95% probability.
    b) Will be between -17.6% to 15.1%, with 95% probability.
    c) Will not change by a statistically significant amount from the first quarter of 2009.
    d) Will be between -11.7% to 4.6%, with 95% probability.
    e) Will be between -9.24% to 7.10%, with 95% probability.

19. Assuming the SRM, the *p*-value for the intercept estimates that if H_0: $\beta_0 = 0$ holds, then the
    a) Estimated intercepts of 12.34 percent of samples lie within 2.31 of 0.
    b) Probability that $\beta_0 = 0$ is 0.1234.
    c) Estimated intercepts of 12.34 percent of samples lie 2.31 or farther from 0.
    d) Probability of a Type II error if we reject H_0 is 0.1234.
    e) Probability of a sample with estimated intercept $b_0 = -2.313$ is 0.1234.

20. The plot at the right shows residuals from the fitted model. From this plot, we should conclude that
    a) The underlying model errors are autocorrelated.
    b) The data have been rounded prior to the analysis.
    c) Prediction intervals from the model are not reliable.
    d) The residuals are consistent with the assumption of normally distributed errors.
    e) The residuals are consistent with the assumption of equal variance.



21. Before relying on inferences from the fitted simple regression model, it is *most* crucial that the day trader
    a) Inspect the leverage plot for this explanatory variable to identify leveraged outliers.
    b) Verify that the variation of the response does not change during this period.
    c) Plot the residuals over time to check for the presence of autocorrelation.
    d) Verify that the response variable is normally distributed.
    e) Obtain longer time series for the explanatory variable and response.

The day trader added a second explanatory variable. The added variable *Quarter* identifies the quarter: Q1, Q2, Q3 or Q4. The following tables summarize the resulting expanded model.

Summary of Fit

RSquare	0.316303
RSquare Adj	0.26268
Root Mean Square Error	7.961069
Mean of Response	1.77082
Observations (or Sum Wgts)	56

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	1495.3863	373.847	5.8986
Error	51	3232.3096	63.379	Prob > F
C. Total	55	4727.6959		0.0006*

Expanded Estimates

Nominal factors expanded to all levels

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2.224099	1.442153	-1.54	0.1292
Lag(Recreation Services % Change)	1.3542097	0.33006	4.10	0.0001*
Quarter[Q1]	1.5901938	1.846861	0.86	0.3933
Quarter[Q2]	-3.53854	1.84278	-1.92	0.0604
Quarter[Q3]	3.184886	1.842631	1.73	0.0900
Quarter[Q4]	-1.236539	1.845547	-0.67	0.5059

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Lag(Recreation Services % Change)	1	1	1066.9143	16.8340	0.0001*
Quarter	3	3	373.5783	1.9648	0.1310

22. The estimated coefficient of *Quarter[Q1]* in the expanded model indicates that when used to describe or predict market movements in the first quarter,
    a) The slope of *Lag*(*Recreation Services % Change*) increases *to* 1.59.
    b) The slope of *Lag*(*Recreation Services % Change*) increases *by* 1.59.
    c) The average percentage change in the market is 1.59.
    d) The estimated intercept to use in this calculation is -0.634.
    e) The estimated intercept to use in this calculation is 1.59.

23. The output shows that the addition of *Quarter* to the simple regression does not produce a statistically significant improvement in the fit of the model because
    a) Each of the slopes associated with *Quarter* is not statistically significant.
    b) The overall *F*-statistic of this model is less than that of the prior simple regression.
    c) The effect test for *Quarter* is not statistically significant.
    d) The increase in R^2 is only 0.08.
    e) The sum of the four estimated slopes for *Quarter* is 0.

24. The best explanation for the fact that the slope of *Lag*(*Recreation Services % Change*) in this regression is nearly identical to its slope in the prior simple regression (1.384) is that
    a) *Quarter* is not statistically significant whereas *Lag*(*Recreation Services % Change*) is.
    b) *Quarter* is a categorical variable whereas *Recreation Services % Change* is numerical.
    c) The sample size is small.
    d) The average of *Recreation Services % Change* is similar in the 4 quarters.
    e) The $R^2 = 0.316$ is relatively small.

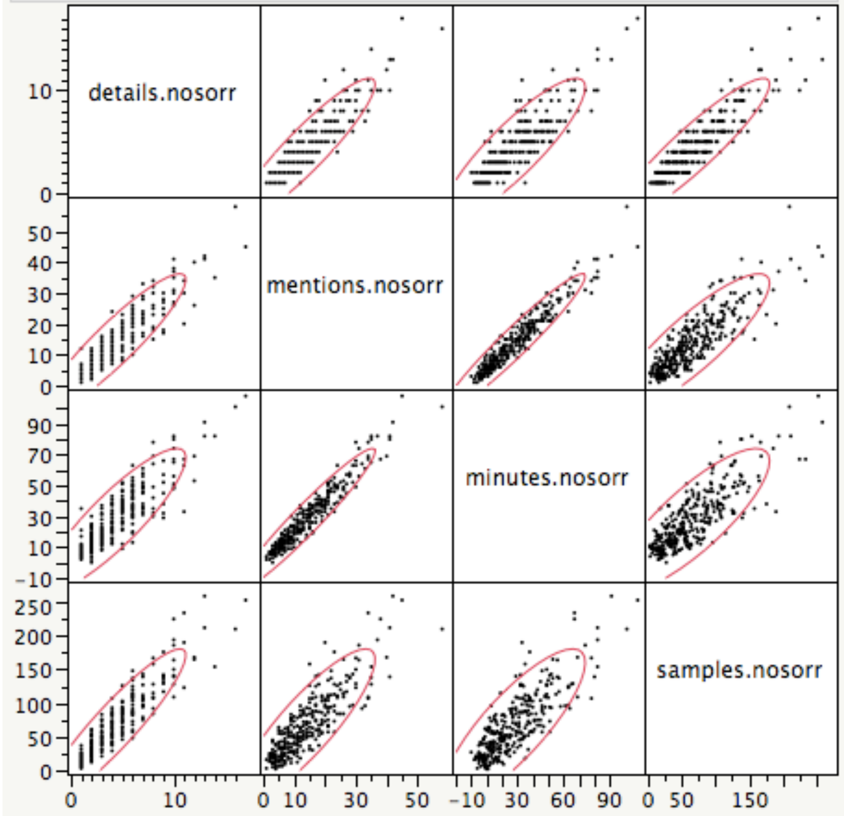
25. The day trader who developed this model suspects that percentage changes in recreational spending have a statistically significantly larger impact on future market returns during Quarter 3 and a smaller impact on future market returns during Quarter 1. In order to investigate this possibility, the day trader should
    a) Add the interaction of *Quarter* and *Lag(Recreational Services % Change)* to the model.
    b) Remove the variables *Quarter[Q2]* and *Quarter[Q4]* from the model.
    c) Fit a simple regression using only the data for Quarter 1 and Quarter 3.
    d) Inspect t-statistics of *Quarter[Q1]* and *Quarter[Q3]* in the expanded model.
    e) Inspect the leverage plot of *Lag(Recreational Services % Change)*.

26. The Durbin-Watson statistic of the expanded model is $D = 2.01$. This indicates that
    a) The underlying model errors are independent.
    b) The underlying model errors are not normally distributed.
    c) The explanatory variables in the expanded model are collinear.
    d) The underlying model errors are not autocorrelated.
    e) The response and explanatory variables are nonlinearly related in the expanded model.

Questions 27-36. Pharmaceutical companies in the US advertise drugs by sending sales representative to visit physicians at their offices. Each visit by a sales representative is known as a "detail". During a detail visit, the sales representative often leaves samples for the physician to use to treat patients. This analysis considers the success of this type of promotion. The response is the number of new prescriptions written by 323 physicians for a new drug (Nosorr) following a promotional campaign. The main competitor of Nosorr is the drug Paenex. The other variables in this model are listed in the table below.

Variable	Definition
Practice size	Number of physicians sharing a group practice
Details Nosorr	Number of detail visits promoting Nosorr
Details Paenex	Number of detail visits by competitor promoting Paenex
Samples Nosorr	Number of samples of Nosorr given to physician
Mentions Nosorr	Number of times Nosorr mentioned during detail visit
Minutes Nosorr	Number of minutes of the Nosorr detail visit
DTC	Amount of direct to consumer advertising in the community of the physician's office (recorded as pages of local newspaper ads)
Years Exper	Years that the physician has been practicing
Specialty	GP: general practitioner PD: pediatrician (children's doctor) RH: rheumatologist (specialist treating arthritis, bone)
Age	Age of the physician, in years

Scatterplot Matrix



Summary of Fit

RSquare	0.716267
RSquare Adj	0.708108
Root Mean Square Error	16.48493
Mean of Response	31.99381
Observations (or Sum Wgts)	323

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	9	214725.33	23858.4	87.7944
Error	313	85058.65	271.8	Prob > F
C. Total	322	299783.99		<.0001*

Expanded Estimates

Nominal factors expanded to all levels

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-20.51032	9.14852	-2.24	0.0257*
practice.size	1.800703	0.508619	3.54	0.0005*
details.nosorr	1.9294091	0.897386	2.15	0.0323*
details.paenex	-2.983819	0.875514	-3.41	0.0007*
samples.nosorr	0.0652567	0.050753	1.29	0.1995
dtc	6.600572	0.492768	13.39	<.0001*
years.exper	-0.690004	0.306121	-2.25	0.0249*
specialty[GP]	-2.548644	1.324073	-1.92	0.0552
specialty[PD]	-21.13215	1.585219	-13.33	<.0001*
specialty[RH]	23.680794	1.630842	14.52	<.0001*
age	0.8925395	0.317797	2.81	0.0053*

27. From the scatterplot matrix on the previous page, we can see that
    a) The highest correlation is between *Minutes* and *Samples*.
    b) The highest correlation is between *Details* and *Samples*.
    c) The average of *Minutes* is larger than 65.
    d) The variable with the largest SD is *Samples*.
    e) The maximum number of *Minutes* is less than 90.

28. The best evidence to show that this multiple regression explains statistically significant variation in the number of new prescriptions written is to observe that
    a) The fitted model explains more than 70% of the variation in new prescriptions.
    b) Numerous explanatory variables have coefficients that are statistically significant.
    c) The RMSE of the fitted model is small compared to the variation of new prescriptions.
    d) Several t-statistics exceed 10 in magnitude.
    e) The p-value of the F-statistic is less than 0.05.

29. Anne and Dave are physicians who are both 40 years old with 12 years of experience and practice in the same community. Both have individual practices (practice size = 1), were detailed 5 times for Nosorr and 4 times for Paenex, and received 10 samples of Nosorr. Anne is a rheumatologist and Dave is a pediatrician. The estimated model
    a) Predicts Anne will write about 23.7 more prescriptions for Nosorr than Dave.
    b) Predicts Dave will write about 21.1 fewer prescriptions for Nosorr than Anne.
    c) Predicts Anne will write about 44.8 more prescriptions for Nosorr than Dave.
    d) Predicts Anne will write the same number of prescriptions for Nosorr as Dave.
    e) Requires that we know the region in order to compare predictions for Anne and Dave.

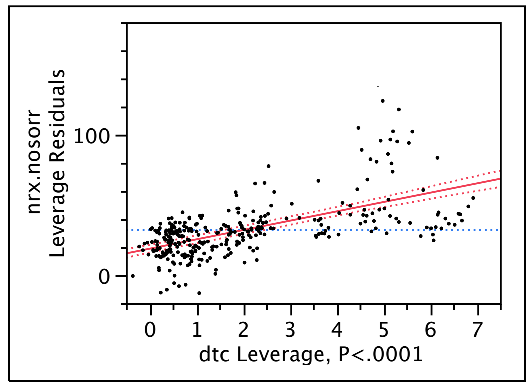
30. According to the estimated model, an increase in DTC from 5 to 10 pages of advertising would, assuming the other types of promotion are held constant, increase the expected number of new Nosorr prescriptions written by a physician by about (assume the MRM)
    a) 32 to 34 more new prescriptions, with 95% confidence.
    b) 28 to 38 more new prescriptions, with 95% confidence.
    c) An amount that can only be determined if the other explanatory variables are given.
    d) 0 to 66 more new prescriptions, with 95% confidence.
    e) 0 to 198 more new prescriptions, with 95% confidence.

31. If a physician receives one detail for Paenex and one for Nosorr, then the combined effect of these details on the average number of new prescriptions for Nosorr (assume the estimated coefficients in the regression are uncorrelated and that the MRM holds)
    a) Changes the average by -3.6 to 1.5 new prescriptions, with 95% confidence.
    b) Changes the average by -1.1 to -4.6 new prescriptions, with 95% confidence.
    c) Changes the average by -2.5 to 4.6 new prescriptions, with 95% confidence.
    d) Increases average use of Nosorr by significantly more than 0 new prescriptions.
    e) Changes the average by -1.4 to -8.5 new prescriptions, with 95% confidence.

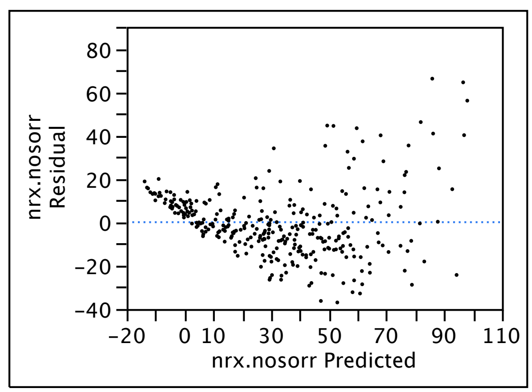
32. If we add the two variables *Mentions Nosorr* and *Minutes Nosorr* to the multiple regression, then we can anticipate from the shown output that
    a) R^2 will increase to more than 0.80.
    b) *RMSE* will remain at 16.485.
    c) The overall *F*-ratio will increase.
    d) The standard error for *Details Nosorr* will increase.
    e) The variable *Samples Nosorr* will become statistically significant.

33. *Years of Experience* has a positive marginal correlation with the number of new prescriptions written for Nosorr, but has a negative coefficient in the shown multiple regression. The best explanation for the change in sign of the direction of the association is
    a) Anything can happen in a multiple regression.
    b) *Years of Experience* is correlated with other explanatory variables.
    c) The marginal correlation is not statistically significant whereas the partial slope is.
    d) The multiple regression omits an important lurking variable.
    e) The multiple regression includes a categorical variable with more than 2 levels.

34. The leverage plot for *DTC* shown to the right of this question indicates that
    a) *DTC* should be transformed to a log scale.
    b) *DTC* is not a statistically significant explanatory variable in the multiple regression.
    c) Some communities receive noticeably higher direct to consumer advertising.
    d) Certain specialties of physicians receive more of this type of advertising than others.
    e) DTC is highly collinear with other explanatory variables in the regression.



35. The scatterplot of residuals on predicted values shown to the right of this question indicates that
    a) The equation under-predicts the use of Nosorr by physicians who write very few prescriptions.
    b) Prediction intervals from this equation are too wide.
    c) Predictions from this equation are more accurate for physicians who write many Nosorr prescriptions.
    d) The equation in general over-predicts the use of Nosorr by physicians.
    e) Predictions for large numbers of Nosorr prescriptions represent extrapolations from the observed data.



36. Having seen the scatterplot of residuals on predicted values (prior question), the most useful next action to take is to
    a) Find the important explanatory variable suggested by the pattern in this plot.
    b) Avoid using prediction intervals because the residuals are not normally distributed.
    c) Transform the response to capture the curvature and changing variation.
    d) Remove extraneous explanatory variables that produce the evident collinearity.
    e) Remove pronounced outlying, highly leveraged observations and refit the equation.

Questions 37-43. Managers at the pharmaceutical firm that makes Nosorr used the data considered in the previous analysis to investigate how sales representatives distribute samples to physicians. They constructed the following regression model with the response *Samples Nosorr*.

Summary of Fit

RSquare	0.849302
RSquare Adj	0.846925
Root Mean Square Error	18.30587
Mean of Response	65.98452
Observations (or Sum Wgts)	323

Expanded Estimates

Nominal factors expanded to all levels

Term	Estimate	Std Error	t Ratio
Intercept	0.9913826	2.23921	0.44
details.nosorr	15.004423	0.541928	27.69
specialty[GP]	-1.66286	3.375637	-0.49
specialty[PD]	3.2978276	2.93553	1.12
specialty[RH]	-1.634968	3.173656	-0.52
specialty[GP]*details.nosorr	1.3862515	0.770312	1.80
specialty[PD]*details.nosorr	-1.664927	0.879518	-1.89
specialty[RH]*details.nosorr	0.2786757	0.628638	0.44

Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
GP	69.052322	1.8043531	68.2330
PD	61.033692	2.8841589	35.3750
RH	64.368730	2.0911797	95.5833

LSMeans Differences Tukey HSD

$\alpha = 0.050$ Q= 2.35485

LSMean[i]	Mean[i]-Mean[j] / Std Err Dif / Lower CL Dif / Upper CL Dif	LSMean[j] GP	PD	RH
GP	Mean[i]-Mean[j]	0	8.01863	4.68359
	Std Err Dif	0	3.40207	2.76201
	Lower CL Dif	0	0.00727	-1.8205
	Upper CL Dif	0	16.03	11.1877
PD	Mean[i]-Mean[j]	-8.0186	0	-3.335
	Std Err Dif	3.40207	0	3.5625
	Lower CL Dif	-16.03	0	-11.724
	Upper CL Dif	-0.0073	0	5.05412
RH	Mean[i]-Mean[j]	-4.6836	3.33504	0
	Std Err Dif	2.76201	3.5625	0
	Lower CL Dif	-11.188	-5.0541	0
	Upper CL Dif	1.82054	11.7242	0

37. The fitted equation predicts that a general practitioner who is visited 10 times by sales representatives (*Details Nosorr* = 10) receives about
    a) 163 samples.
    b) 81 samples.
    c) 151 samples.
    d) 12 samples.
    e) 15 samples.

38. The fitted equation indicates that general practitioners receive on average about
    a) 1.4 samples per detail visit.
    b) 16.4 samples per detail visit.
    c) 13.3 samples per detail visit.
    d) 15.0 samples per detail visit.
    e) 2.4 samples per detail visit.

39. At the average level of detailing (4.25 details overall) general practitioners receive on average about
    a) 33 more samples than pediatricians.
    b) 3.3 more samples than pediatricians.
    c) 8.0 more samples than pediatricians.
    d) 4.7 more samples than pediatricians.
    e) 5.0 fewer samples than pediatricians.

40. When detailed the average number of visits (4.25), general practitioners receive on average (assuming that the MRM holds)
    a) Statistically significantly more samples than pediatricians.
    b) Statistically significantly more samples than pediatricians and rheumatologists.
    c) Statistically significantly fewer samples than pediatricians.
    d) Statistically significantly fewer samples than pediatricians and rheumatologists.
    e) About the same number of samples as pediatricians and rheumatologists.

41. It has been claimed that sales representatives on average give more samples per detail visit to some types of physicians than to others. In order to test whether such differences are statistically significant requires (assume the MRM holds)
    a) Adding the interaction of samples and specialities to the regression model.
    b) Inspecting side by side boxplots of samples grouped by specialty of physician.
    c) Two-sample *t*-tests comparing the level of sampling in the three types of specialties.
    d) Checking the effect test output for the interaction of specialty and detailing.
    e) That sales representatives visit physicians of each specialty an equal number of times.

42. The most important diagnostic plot that should be considered next before accepting the use of the MRM with these data is
    a) A sequence plot of the residuals to check for the presence of autocorrelation.
    b) The leverage plot for *Details*.
    c) Comparison (side-by-side) boxplots of Samples Nosorr grouped by specialty.
    d) Comparison (side-by-side) boxplots of the residuals grouped by specialty.
    e) A normal quantile plot of the residuals.

43. A carton of samples contains 6 individual sample packages. If the response variable in this analysis is changed from the number of sample packages to the number of cartons, then
    a) R^2, RMSE and *t*-statistics remain the same as shown.
    b) Standard errors of all estimated coefficients become 6 times larger.
    c) All estimated slopes that involve detailing become 6 times larger.
    d) All of the *t*-statistics remain as shown.
    e) R^2 and RMSE remain as shown.