Statistics 608-621 Waiver Exam

August 24, 2004

On the answer sheet…

– Use a **#2 pencil**. Erase changes completely.

– **Fill in your name and student id number**.

– **Mark the "bubbles"** under the letters of your name and student id number on the form. Failure to do so will lead to a score of zero.

– Choose the **one *best* answer** for each question.

Turn in the solution page only; keep the test. Mark your choices on your copy of the exam. The solutions will be posted in WebCafé in several days, and you can use the "Gradebook" feature to check your results.

You may consult **1 textbook and 2 pages of notes** during the exam. No laptops or other computers are allowed. No other reference materials are permitted.

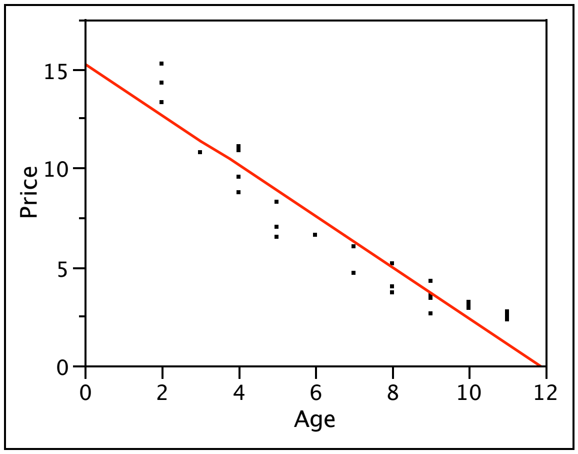
You have **two hours** for the exam. The **computer output** associated with one or more items should be considered an essential part of the questions.

The multiple-choice questions are equally weighted. Your score is the number of correct answers given to questions that are scored. Some questions may be dropped and not counted as part of the overall score. There is no deduction for incorrect answers.

STOP

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

(Questions 1–12) A student interested in purchasing a used Honda Accord collected the asking price for 4-door, 4-cylinder used Accords from a newspaper. The paper listed the price of 30 comparable models of various ages (in years). A new car has age 0; a car that has been used for 2 years has age 2. The price is listed in thousands of dollars.

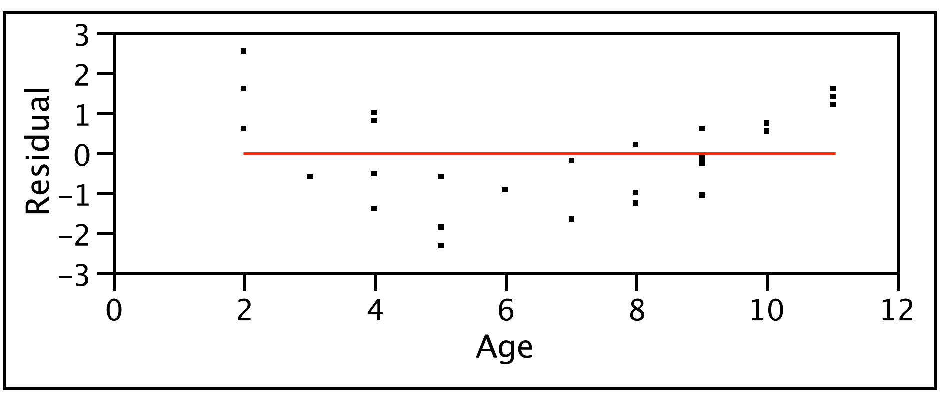


Summary of Fit

RSquare	0.913
Root Mean Square Error	1.215
Mean of Response	6.610
Observations (or Sum Wgts)	30

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	15.29	0.55	27.69	<.0001
Age	−1.29	0.08	−17.17	<.0001



(1) The fitted model implies that the expected price of a brand new Honda Accord is
(a) About \$6,610.
(b) About \$9,130.
(c) About \$12,900.
(d) About \$15,300.
(e) More than \$20,000.

(2) According to the fitted model, the asking price for a 2000 Honda Accord on average is
(a) About \$2,600 more than the price of a 2002 Accord.
(b) About \$1,300 more than the price of a 2002 Accord.
(c) About the same as the price of a 2002 Accord.
(d) About \$1,300 less than the price of a 2002 Accord.
(e) About \$2,600 less than the price of a 2002 Accord.

(3) This model predicts the asking price for a 10-year old used Honda Accord to be
(a) \$15,290
(b) \$6,610
(c) \$2,390
(d) \$1,529
(e) \$1,290

(4) A buyer with \$5,000 seeks a used Accord. What is the chance of the buyer having enough money to purchase an 8-year-old Accord? (Use the fitted model and assume the usual regression conditions.)
(a) 16%
(b) 33%
(c) 50%
(d) 84%
(e) 95%

(5) A different newspaper contained a larger listing of 60 used Accords of similar ages. Had this model been fit to the larger sample then we can be sure that
(a) The R^2 of the model would have been larger.
(b) The RMSE of the model would have been smaller.
(c) The standard error of the slope would have been smaller.
(d) The standard deviation of the residuals would have been smaller.
(e) All of the above would have occurred.

(6) The plot of the residuals from this model suggests that
(a) The fitted model should be non-linear.
(b) The errors underlying the model are autocorrelated.
(c) The errors underlying the model lack constant variance.
(d) The errors underlying the model are not normally distributed.
(e) The fitted model meets the usual assumptions.

(7) A national study by Honda claimed that five-year-old Accords depreciate about $1,400 during the year. Based on this regression, the claimed level of depreciation is
(a) Larger than that of the fitted model, but not by a statistically significant margin.
(b) Larger than that of the fitted model, by a statistically significant margin.
(c) Essentially the same as the value suggested by this model.
(d) Much smaller (closer to zero) than the value implied by this model.
(e) The fitted model does not offer a value for comparison.

(8) A friend of the buyer playing with the JMP modeling software accidentally told the software to treat the variable *Age* as categorical (rather than continuous). If *Price* is regressed on *Age* as a categorical variable then
(a) The slope in the alternative model will be steeper.
(b) The intercept in the alternative model will be smaller.
(c) The RMSE of the alternative model will be larger.
(d) The R^2 of the alternative model will be larger.
(e) The software will indicate an error and not produce an alternative model.

(9) Which of the following additional data would produce the most precise estimate in this model of the effects of aging on the asking price of a used Honda Accord?
(a) Adding the prices and ages of 10 one-year-old cars.
(b) Adding the prices and ages of 10 three-year-old cars.
(c) Adding the prices and ages of 10 six-year-old cars.
(d) Adding the prices and ages of 10 nine-year-old cars.
(e) All of these additions would be expected to offer the same improvement.

A friend of the student who is majoring in economics fit the following different model to prices and ages of these 30 cars. She regressed the log of the price on the log of the age of the car, and obtained the following results. She used logs to base *e*.

Summary of Fit

RSquare	0.927
Root Mean Square Error	0.165

Parameter Estimates

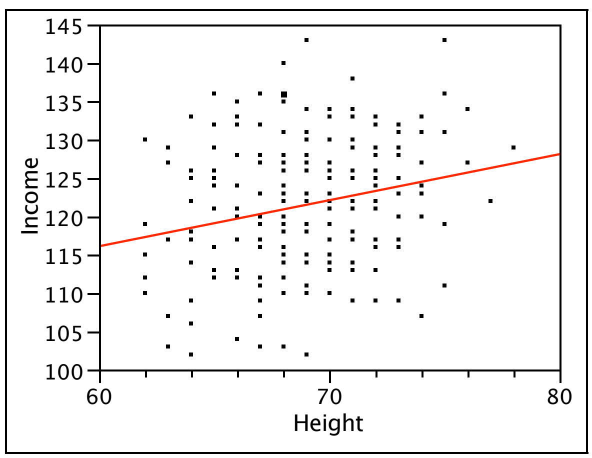
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3.536	0.101	34.91	<.0001
Log(Age)	–1.025	0.054	–18.84	<.0001

(10) Compared to the initial linear model, her model using logs
(a) Explains statistically significantly more variation.
(b) Explains more variation, but not by a statistically significant margin.
(c) Corrects for problems caused by autocorrelation in the data.
(d) Captures nonlinear patterns missed by the linear model.
(e) Cannot be interpreted because of the use of logs for the response.

(11) Based on the shown summary of her log-log model,
- (a) A new Accord costs about $35,400.
- (b) Predicted prices are within $330 of asking prices with 95% confidence.
- (c) The price of a used Accord drops about $1000 for each additional year of aging.
- (d) The price of a used Accord drops about 1% for each 1% increase in age.
- (e) The price of a used Accord drops about 1% for each additional year of age.

(12) Using the model with logs and accepting the usual assumptions, an asking price of $12,000 for a 4 year old Accord is
- (a) Much higher than predicted by this model, suggesting the car is over-priced.
- (b) Higher than predicted by this model, but within typical variation.
- (c) Close to the predicted value from this model.
- (d) Lower than predicted by this model, but within typical variation.
- (e) Much lower than predicted by this model, suggesting the car is a bargain.

(Questions 13–20) Several years ago, a magazine printed an article that claimed taller executives were higher paid than their shorter counterparts. A rival magazine gathered a sample to see whether this claim was plausible. The rival collected a survey of 250 mid-level executives with comparable experience and qualifications. The shown fitted model regresses the reported income (in thousands of dollars) on the height (in inches).



Summary of Fit

RSquare	0.05
Root Mean Square Error	8.3
Mean of Response	121.6
Observations (or Sum Wgts)	250

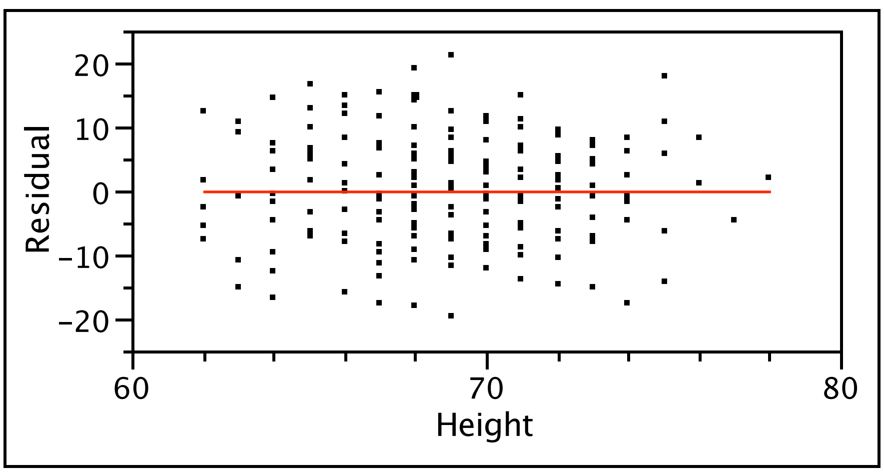
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	79.93	11.48	6.97	<.0001
Height	0.60	0.17	3.63	0.0003

(13) The results show that, given the usual assumptions of linear regression, that *Height*
(a) Has statistically significant positive association with *Income*.
(b) Has a small, insignificant relationship with *Income*.
(c) and *Income* are independent.
(d) and *Income* are uncorrelated.
(e) and *Income* are Normally distributed.

(14) Based on the fitted model and usual assumptions, with 95% confidence, a manager who is 75 inches tall would make on average how much more than a manager who is 65 inches tall?
(a) Between \$4,340 and \$7,660 more.
(b) Between \$2,600 and \$9,400 more.
(c) Between \$4,300 and \$7,700 more.
(d) Between \$5,966 and \$6,034 more.
(e) They would earn about the same income.

(15) Based on the fitted model and usual assumptions, you can predict that 2/3 of the managers who are 70 inches tall make between
(a) \$104,930 to \$138,930
(b) \$110,000 to \$135,000
(c) \$105,330 to \$138,530
(d) \$113,630 to \$130,230.
(e) \$121,760 to \$122,100.



(16) The residual plot (shown above) for this model implies that
(a) The data are not independent.
(b) The errors are autocorrelated.
(c) The errors are heteroscedastic.
(d) The errors are not normally distributed.
(e) The heights have been rounded to inches.

(17) The p-value shown for the slope of *Height* in the fitted model implies, given the usual assumptions, that
(a) There is a 3% chance that the true slope for *Height* is zero.
(b) There is a 0.03% chance that the true slope for *Height* is zero.
(c) There is a 0.03% chance that the true slope for *Height* is 0.60.
(d) The true slope for *Height* lies within 2×0.0003 of 0.60 with probability 0.95.
(e) It is unlikely to estimate a slope for *Height* this large if the true slope is zero.

(18) If the heights of the managers had been measured in centimeters rather than inches, then (2.54 cm = 1 inch)
(a) The R^2 of the model would have been larger.
(b) The RMSE of the fitted model would have been smaller.
(c) The RMSE of the fitted model would have been larger.
(d) The value of the fitted slope would be about 0.236.
(e) The value of the fitted slope would be about 1.524.

(19) It was later discovered that some of these managers were men, and some were women. To investigate whether height has a larger effect on income for women than for men, we should
(a) Add a lagged variable for last year's height.
(b) Check the assumption of equal variance using boxplots grouped by sex.
(c) Add a categorical variable indicating the sex of the manager.
(d) Add the interaction between a categorical variable for sex with height.
(e) Do both "c" and "d".

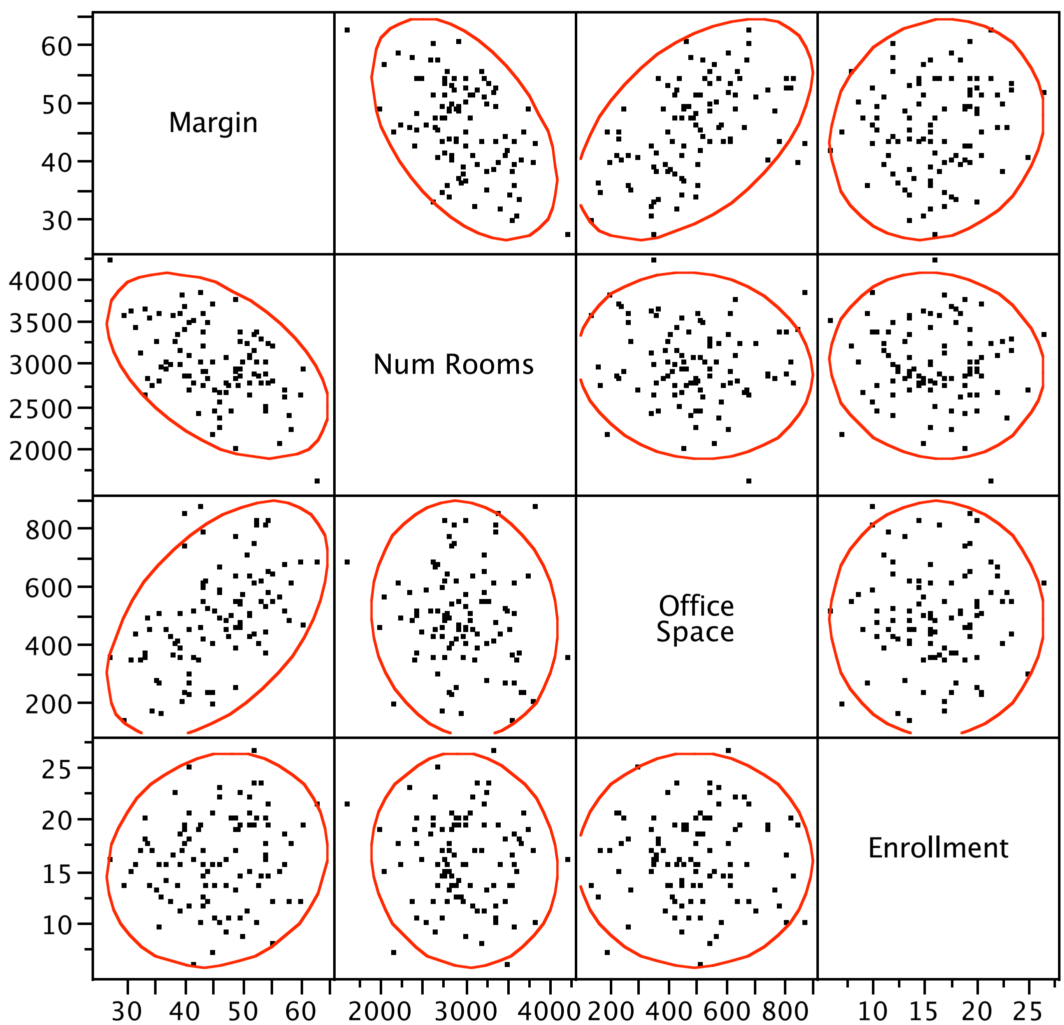
(20) In order to produce a statistically significant improvement in the fit of this model, one added variable would have to increase the shown R^2 to be at least (that is, choose the smallest value that would produce a significant improvement)
(a) 7.5%
(b) 10%
(c) 15%
(d) 25%
(e) 50%

(Questions 21–28) A national motel chain has a model for the operating margin of its franchises. The operating margin is defined to be the ratio of net profit to total revenue (as a percentage). The company plans to use this model to help it identify new profitable sites. The following analysis considers data from 100 motels operated by this chain. The variables are

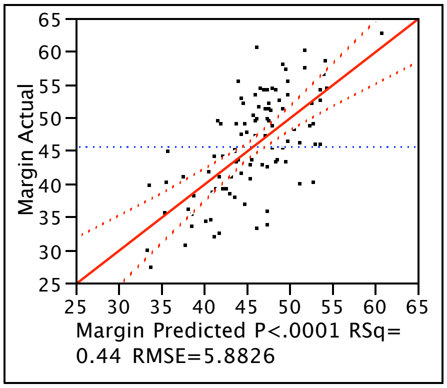
Margin	100 times the ratio of net profit to revenue.
Num rooms	Number of hotel rooms and motel rooms within 3 miles of the site.
Office space	Number of square feet (in 1000's) of surrounding office.
Enrollment	Number enrolled (in 1000's) in nearby colleges and universities.

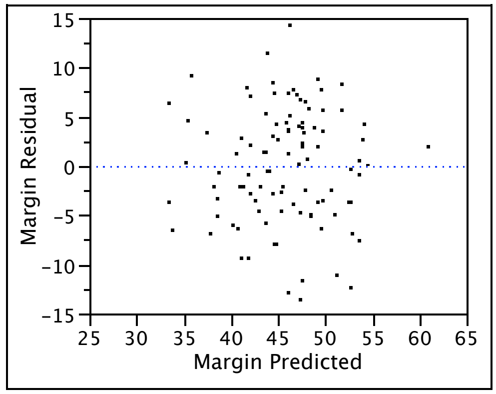
Correlations

	Margin	Num Rooms	Office Space	Enrollment
Margin	1.000	-0.470	0.500	0.123
Num Rooms	-0.470	1.000	-0.093	-0.064
Office Space	0.500	-0.093	1.000	-0.001
Enrollment	0.123	-0.064	-0.001	1.000



(21) The scatterplot matrix and correlations among these variables imply that
(a) A regression of *Margin* on these 3 predictors explains more than ½ of the variation.
(b) More students are enrolled in locations with more office space.
(c) No one predictor can explain more than ¼ of the variation in *Margin*.
(d) Outliers will distort the multiple regression of *Margin* on these 3 predictors.
(e) High collinearity will complicate the interpretation of a multiple regression for *Margin* on these predictors.





Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	2627.40	875.801	25.3087
Error	96	3322.06	34.605	**Prob > F**
C. Total	99	5949.46		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	53.9826	5.1777	10.43	<.0001
Num Rooms	-0.0073	0.0013	-5.48	<.0001
Office Space	0.0216	0.0036	6.03	<.0001
Enrollment	0.1776	0.1406	1.26	0.2094

(22) Two motels operated by this chain reside in similar communities, with the exception that one is located near a business office complex with 300,000 square feet of office space. The other is not near any offices. We should expect (based on this model) the average operating margin of the motel near the business complex to be
(a) About the same as the margin of the motel that is not near offices.
(b) 6½% higher than the margin of the motel that is not near offices.
(c) 54%.
(d) 2.2% higher than the margin of the motel that is not near the offices.
(e) Near 100%.

(23) A consultant offered the management of the motel chain an accounting model fit to the same data that she claims captures 35% of the variation in operating margins. The results for the shown model imply that

(a) Predictions from her model will be *more* accurate than those from this model.
(b) Predictions from her model have comparable accuracy to those from this model.
(c) Predictions from her model will be *less* accurate than those from this model.
(d) Her model has been over-fit to the data and is not reliable.
(e) The shown model has omitted important predictors that should be added to improve the fit.

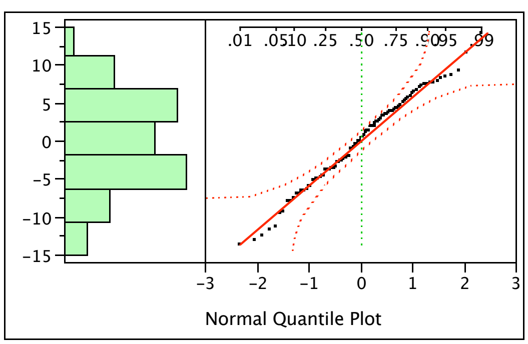
(24) The shown summary of this regression model implies that this model

(a) Explains a statistically significant proportion of the variation in *Margin*.
(b) Does not explain a statistically significant proportion of the variation in *Margin*.
(c) Should be fit to a smaller sample to save degrees of freedom in testing.
(d) Requires a larger sample size to obtain a statistically significant fit.
(e) Explains more than half of the variation in operating margin.

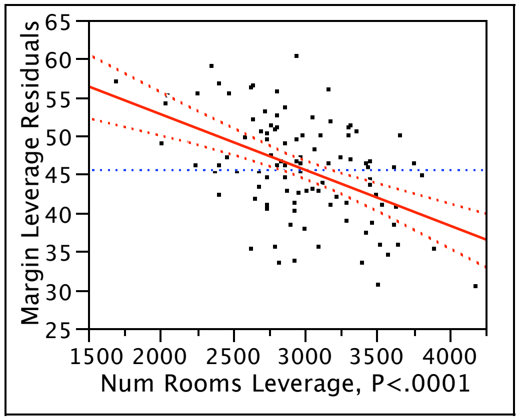
(25) If we use this model to predict the operating margin of a motel in a location with 255,200 square feet of nearby office space, 22,600 students enrolled in a nearby college and 2500 competing motel rooms, then

(a) With 95% confidence, the predicted margin would be accurate to ±6%.
(b) With 95% confidence, the predicted margin would be accurate to ±12%.
(c) With 95% confidence, the predicted margin would be accurate to ±24%.
(d) The prediction would be unreliable due to the lack of constant variation.
(e) The prediction would be unreliable because we were extrapolating too far from the data used to construct the model.

(26) The following plot of the residuals from this model implies that



(a) The model omits important predictors.
(b) The errors of the underlying model are dependent.
(c) The errors of the underlying model lack constant variance.
(d) The errors of the underlying model have constant variance.
(e) The errors of the underlying model appear normally distributed.



(27) The leverage plot for the effect of *Num Rooms* in the regression (shown above) implies that

(a) Leveraged outliers have increased the standard error of the slope for this predictor.
(b) Leveraged outliers have decreased the slope for this predictor.
(c) Collinearity has increased the standard error of the slope for this predictor.
(d) The motel chain should build more motels with about 1500 to 2000 rooms.
(e) There appear to be no problems in the estimation of this slope.

(28) To obtain a model that is capable of producing substantially more accurate predictions of operating margin, the motel chain is most likely to achieve the greatest success by

(a) Increasing the sample size and refitting the current model.
(b) Transforming the variables to capture the evident increasing returns to scale.
(c) Adjusting for time series factors that have been ignored.
(d) Adding other predictors such as the distance to the nearest competitor.
(e) Removing *Enrollment* from this model.

Questions (29-34) A resort hotel in Jamaica would like to predict the level of occupancy in coming seasons in order to better schedule its staff. To help the resort, a consultant used quarterly occupancy rates during the 5 years 1999-2003 to build the following model. The response in this model is *Occupancy*, defined as the ratio (number of occupied rooms)/(number of available rooms). The variable *Year* is coded in order as 1999, 1999, 1999, 1999, 2000, 2000, 2000, 2000, 2001, …, 2003 and the variable *Quarter* is coded as "Q1", "Q2", "Q3", and "Q4". The data was put in the JMP spreadsheet in time order.

RSquare	0.892
Root Mean Square Error	0.038
Mean of Response	0.695
Observations	20

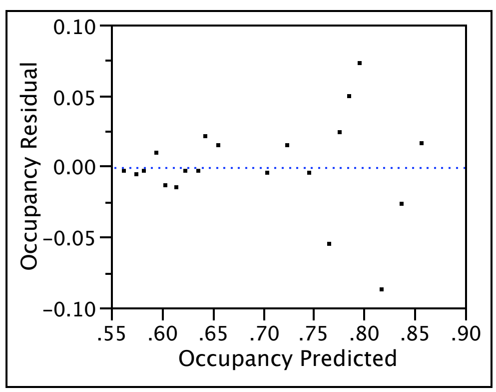
Parameter Estimates

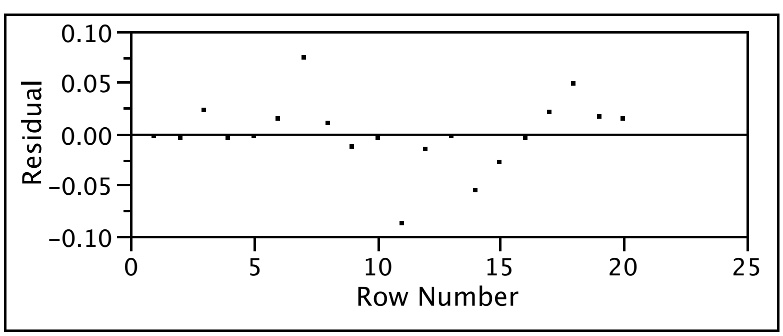
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-39.415	11.953	-3.30	0.0049
Quarter[Q1]	-0.092	0.015	-6.28	<.0001
Quarter[Q2]	0.023	0.015	1.49	0.1582
Quarter[Q3]	0.122	0.015	8.30	<.0001
Quarter[Q4]	-0.080	0.015	-5.43	<.0001
Year	0.020	0.006	3.36	0.0043

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Quarter	3	3	0.1604	37.355	<.0001
Year	1	1	0.0161	11.260	0.0043

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
1.3278658	20	0.3310	0.0715





(29) Based on the fit of this model, the predicted level of occupancy at the resort for the first quarter of 2004 is approximately
(a) Less than 50%
(b) 60%
(c) 70%
(d) 80%
(e) 100%

(30) Accepting the model with the usual assumptions, we can conclude that occupancy rates at this resort in Quarter 4 are, on average,
(a) About the same as the rate in the prior Quarter 3.
(b) Significantly lower than the rate in the prior Quarter 3.
(c) Significantly higher than the rate in the prior Quarter 3.
(d) About 31%.
(e) Unchanged over these 5 years of data.

(31) The resort used a new advertising campaign that promised to raise occupancy rates in a future quarter by 5% (e.g., from 80% to 85%). If we accept the usual assumptions, can this model be used to judge whether the campaign achieved its goal?
(a) No.
(b) Yes.
(c) Yes, but only if we know the quarter being predicted.
(d) Yes, but only if the campaign achieves a higher gain (say 10%) than promised.
(e) We cannot answer this question without knowing the predicted occupancy rate.

(32) It has been conjectured that the rate of occupancy has been growing slower during the 4^{th} quarter than other quarters. According to this model (and accepting the usual conditions), we can see that
(a) Growth is slower in Q4 than Q3, but not by a significant margin.
(b) Growth is slower in Q4 than Q3, by a statistically significant margin.
(c) A longer time series of data would be necessary to address this conjecture.
(d) Equivalent standard errors for the quarter effects show that growth is similar in all.
(e) The model does not address this conjecture.

(33) During this time period, the occupancy rates of resorts like this one in Jamaica grew by 3% per year. Compared to this rate, the fit of this model (along with the usual assumptions) implies that the occupancy rate at this resort is
(a) Statistically significantly lower.
(b) Lower, but not by a significant margin.
(c) About the same.
(d) Higher, but not by a significant margin.
(e) Statistically significantly higher.

(34) A useful next step in the use of this model would be to
(a) Remove the insignificant indicator for Q2.
(b) Use a partial-F test to assess the overall effect of differences in quarters.
(c) Convert the two predictors *Quarter* and *Year* into a simple time sequence.
(d) Discover why results for Q3 are so variable over these years.
(e) Use any of the above actions to improve the model.

(Questions 35-45) Absenteeism is a problem for some companies. Two economists collected data from a sample of 100 companies in order to explore how various characteristics of these companies affect the rate of absenteeism. The response variable in the following models is *Absent*, the *average number of days that employees are absent* during the year (i.e., total number of days missed due to absence divided by the number of full time equivalent employees). The predictors considered in this analysis are

Wage	Average employee wage, in 1,000's of dollars.
Union %	Percentage of the work force in a union (0 – 100).
Part-time %	Percentage of the work force who work part time (0 – 100).
Union relations	Indicator of good/bad relationship between the union and mgmt
Shift work avail	Indicator of whether shift work is available.

Shift work allows employees to alter their schedule to different times during the week rather than work a fixed schedule, such as 9-5 during the week. The following results summarize the fit of a multiple regression model.

Summary of Fit

RSquare	0.445
Root Mean Square Error	2.618
Mean of Response	6.332
Observations (or Sum Wgts)	100.000

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	517.5928	103.519	15.1008
Error	94	644.3848	6.855	**Prob > F**
C. Total	99	1161.9776		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	12.82	2.08	6.17	<.0001
Wage	−0.16	0.04	−4.05	0.0001
Part-time %	−0.12	0.03	−3.56	0.0006
Union %	0.06	0.01	4.62	<.0001
Shift Work Avail[No]	−0.42	0.28	−1.51	0.1425
Shift Work Avail[Yes]	0.42	0.28	1.51	0.1425
Union Relations[Good]	−1.23	0.27	−4.50	<.0001
Union Relations[Poor]	1.23	0.27	4.50	<.0001

(35) Suppose that two companies both pay an average wage of $50,000 with 35% part-time workers, 50% union participation, and no shift work available. If one company has good relations with the employee union whereas the other has poor relations, then this model predicts that absenteeism at the company with good relations will be

(a) Lower by about 1.23 days per year.
(b) Lower by about 2.46 days per year.
(c) Lower by less than 1 day per year.
(d) Higher by about 1.23 days per year.
(e) Higher by about 2.46 days per year.

(36) If the predictor *Union%* were removed from this model, then

(a) The *RMSE* of the model would increase by a statistically significant amount.
(b) The *RMSE* of the model would decrease, but not by a statistically significant amount.
(c) The R^2 of the model would drop by less than 0.01%.
(d) The slopes of the other predictors would be unchanged.
(e) The F-ratio of the model would be unchanged.

(37) Based on this model, do companies offering shift work have more absenteeism than comparable companies that do not offer shift work?

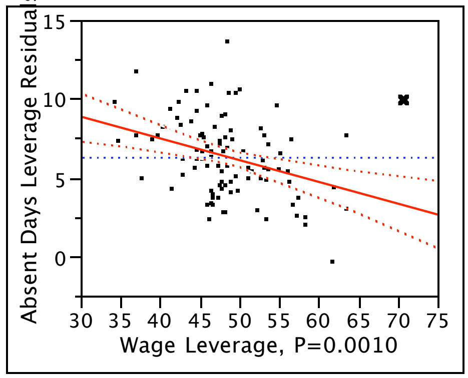
(a) Yes, by a statistically significant amount.
(b) Yes, but not by a statistically significant amount.
(c) No, they have statistically significantly lower absenteeism.
(d) No, they have lower absenteeism.
(e) The effect of union relations is too small to be meaningful.

(38) If an interaction between *Shift Work* and *Union Relations* was added to the model, then the model could answer which of the following questions that it cannot as shown?

(a) Is more variation in absenteeism explained by *Shift Work* or *Union Relations*?
(b) Does the effect on absenteeism of having shift work available depend on the quality of union relations?
(c) Does the elasticity of absenteeism with respect to the percentage in the union depend upon the quality of union relations?
(d) Do *Shift Work* and *Union Relations* jointly explain significant variation in absenteeism?
(e) Has collinearity affected the accuracy of the slopes of *Shift Work* or *Union Relations*?

(39) In order to check the assumption of equal error variation, which of the following plots should be viewed?

(a) A plot of absenteeism on the predicted values from the model.
(b) A plot of absenteeism on each predictor.
(c) The scatterplot matrix of all of the variables in this analysis.
(d) Comparison boxplots of the residuals grouped by *Shift Work* or *Union Relations*.
(e) A normal quantile plot of the residuals.



(40) If the point market with an "*x*" in the shown leverage plot were excluded from this analysis, then

(a) The estimated slope for wage would no longer be significant.
(b) The estimated slope for wage would become more negative.
(c) The estimated slope for wage would become closer to zero.
(d) The RMSE of the fitted model would increase.
(e) The standard error for the slope of *Wage* would increase.

The previous fitted model was expanded by adding two additional predictors. A summary of the *expanded model* follows.

Summary of Fit

RSquare	0.522
Root Mean Square Error	2.456

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	606.9039	86.7006	14.3701
Error	92	555.0737	6.0334	**Prob > F**
C. Total	99	1161.9776		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prb>|t|
Intercept	11.97	2.06	5.80	<.0001
Wage	-0.14	0.04	-3.41	0.0010
Part-time %	-0.12	0.03	-3.65	0.0004
Union %	0.06	0.01	4.38	<.0001
Shift Work Available[No]	-0.60	0.26	-2.28	0.0248
Shift Work Available[Yes]	0.60	0.26	2.28	0.0248
Union Relations[Good]	-1.27	0.26	-4.95	<.0001
Union Relations[Poor]	1.27	0.26	4.95	<.0001
(Wage-48)*Shift Work Avail[No]	0.11	0.04	2.60	0.0109
(Wage-48)*Shift Work Avail[Yes]	-0.11	0.04	-2.60	0.0109
(Wage-48)*Union Relations[Good]	0.09	0.04	2.43	0.0172
(Wage-48)*Union Relations[Poor]	-0.09	0.04	-2.43	0.0172

(41) The fit of this model implies that if a company with average wages $40,000 increases the average wage to $41,000 without affecting its other characteristics (50% part-time, 40% union, no shift work, good union relations), then it can expect that the average level of absenteeism will

(a) Decrease by 1.4 days per year.
(b) Decrease by 0.14 days per year.
(c) Decrease by 0.25 days per year.
(d) Decrease by 0.34 days per year.
(e) Increase by 0.06 days per year.

(42) The fit of the expanded model

(a) Explains about the same amount of variation in absenteeism as the first model.
(b) Significantly more variation in absenteeism than the first model.
(c) Cannot be directly compared to that of the original model without Effect Tests.
(d) Suffers from over-fitting due to the addition of so many more predictors.
(e) Is inferior to the first model because the overall F-ratio has decreased.

(43) The slope for the predictor *Wage* changes from -0.16 in the original model to -0.14 in the expanded model because

(a) The added predictors are correlated with *Wage*.
(b) The added predictors are interactions.
(c) The added predictors are correlated with both *Wage* and *Absent*.
(d) The expanded model has a larger R^2 than the original model.
(e) The added predictors introduce outliers that alter the slopes of the original model.

(44) The coefficient of (Wage-48)*Union Relations[Good] implies that, on average in companies with similar other characteristics and some shift work available, those with good union relationships

(a) Have statistically significantly lower rates of absenteeism.
(b) Have about the same rates of absenteeism as those with poor relationships.
(c) Achieve smaller decreases in absenteeism as wages increase.
(d) Pay about the same wages as those with poor relationships.
(e) Pay statistically significantly higher wages than those with poor relationships.

(45) A company with 100 employees has an average wage of $50,000 with 35% part-time workers, 50% union participation, good union relations, and no shift work available. Given the usual regression assumptions, the **total** number of absent days in the next near is predicted to fall

(a) Between 0 to 720 days, with 95% probability.
(b) Between 225 to 235 days, with 95% probability.
(c) Between 0 and 475 days, with 95% probability.
(d) Between 0 and 877 days, with 95% probability.
(e) Fewer than 100 days, with 95% probability.