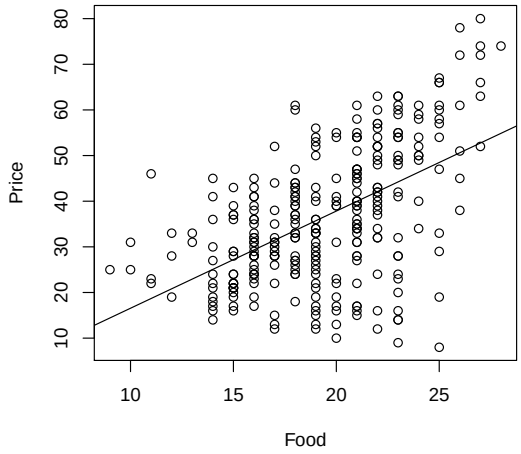
Regression Inference
STAT-UB.0003: Regression and Forecasting Models

Inference

1. Recall the restaurant data: 294 New York City restaurant's from the 2003 Zagat guide. Here is a scatterplot of the data, along with the least squares regression fit:



Here is the Minitab regression output:

```
Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
12.5559  27.93%     27.68%      26.86%

Coefficients

Term       Coef  SE Coef  T-Value  P-Value   VIF
Constant  -4.74     3.95    -1.20    0.232
Food      2.129    0.200    10.64    0.000  1.00

Regression Equation

Price = -4.74 + 2.129 Food
```

(a) What is a reasonable population to go along with this sample?

(b) What is the difference between the true regression parameters (β_0 and β_1) and the regression estimates ($\hat{\beta}_0$ and $\hat{\beta}_1$)?

(c) Construct a 95% confidence interval for β_1, the coefficient of "Food".

(d) What is the meaning of the confidence interval for β_1?

(e) What is the meaning of a 95% confidence interval for β_0? Is this useful for the restaurant example?

(f) Perform a hypothesis test at level 5% of whether or not there is a linear relationship between Price and Food.

2. We used the prices and sizes of 18 apartments in Greenwich Village to fit the model

$$\text{Price} = \beta_0 + \beta_1 \, \text{Size} + \varepsilon,$$

where price is measured in units of \$1000 and size is measured in units of 100 ft^2.

```
Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
101.375  86.87%     86.05%      81.13%

Coefficients

Term       Coef  SE Coef  T-Value  P-Value   VIF
Constant  182.3     62.4     2.92    0.010
Size      44.95     4.37    10.29    0.000  1.00

Regression Equation

Price = 182.3 + 44.95 Size
```

(a) What is a reasonable population for this sample?

(b) Construct a 95% confidence interval for β_1.

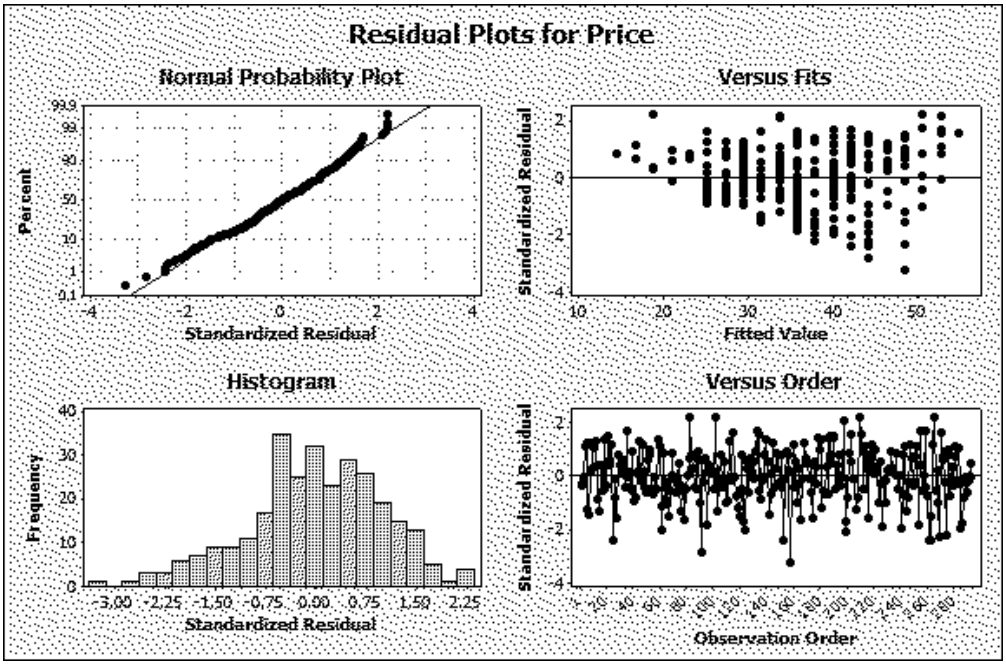
(c) What is the meaning of the confidence interval for β_1?

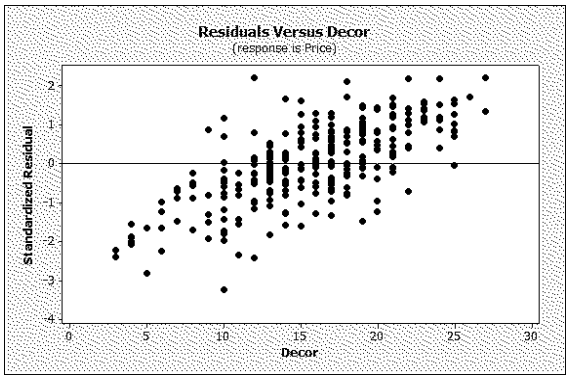
(d) What is the meaning of a 95% confidence interval for β_0? In the context of the housing data, is this useful?

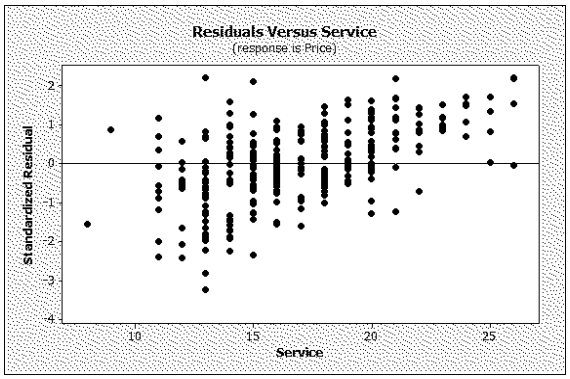
(e) Perform a hypothesis test at level 5% of whether or not there is a linear relationship between Size and Price.

Model assumptions

3. Here are some plots of the residuals from the fit of Price to Food for the Zagat data:







Use the plots to assess whether or not the four regression assumptions hold.