Covariance (Review)

1. Suppose X and Y are random variables with $\text{var}(X) = 4$, $\text{var}(Y) = 3$, and $\text{cov}(X, Y) = -2$.

   (a) Find $\text{cov}(X, 2X + Y)$.

   (b) Find $\text{cov}(4X - Y, Y)$.

   (c) Find $\text{cov}(2X + Y, X + 3Y)$.

Portfolio Optimization (Review)

2. Suppose there are two stocks, X and Y. The annual returns for these stocks can be modeled as *dependent* random variables with correlation ρ. Suppose that the expected returns for the two stocks are both equal to 5%, but the standard deviations for the two stocks are different. Specifically, the standard deviations for stocks X and Y are 1% and 2%, respectively. Consider two strategies for investing $100 between the two stocks:

$$A = 100X,$$
$$B = 50X + 50Y.$$

When does strategy B have more risk than strategy A?

Linear Regression

3. In the following scenarios, which would you consider to be predictor (x) and which would you consider to be response (y)?

   (a) Sales revenue; Advertising expenditures
   (b) Starting salary after college; Undergraduate GPA
   (c) The current month's sales; the previous month's sales
   (d) The size of an apartment; the sale price of an apartment.
   (e) A restaurant's Zagat Price rating; a restaurant's Zagat Food rating.

4. Let y be the payment (in dollars) for a repair which takes x hours. Suppose that
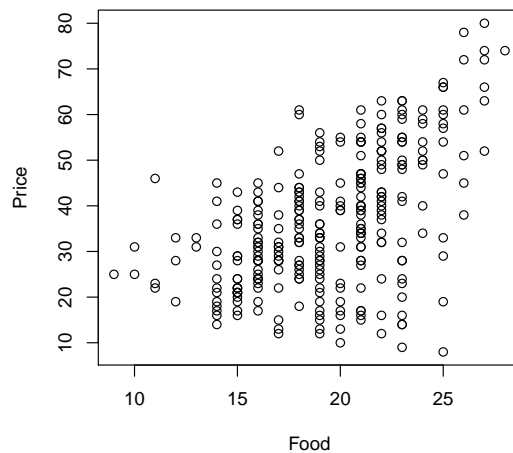
$$y = 25 + 30x.$$

What is the interpretation of this model?

5. Consider two variables measured on 294 restaurants in the 2003 Zagat guide:

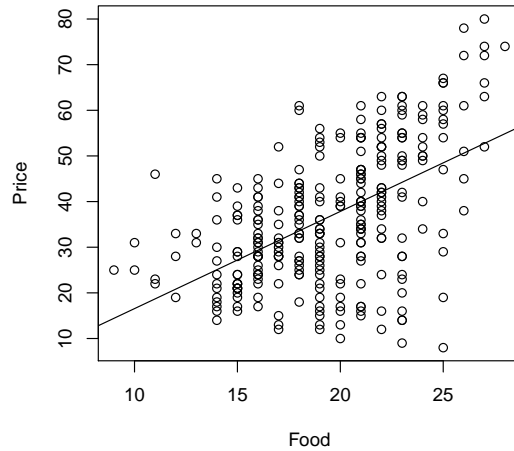$$y = \text{typical dinner price, including one drink and tip (\$)}$$
$$x = \text{Zagat quality rating (0--30).}$$

Here is a scatterplot of y on x:

Why is an exact linear relationship inappropriate to describe the relationship between y and x?

6. Here is the least squares regression fit to the Zagat restaurant data:

Here is the Minitab output from the fit:

```
The regression equation is
Price = - 4.74 + 2.13 Food


Predictor    Coef  SE Coef      T      P
Constant   -4.736    3.950  -1.20  0.232
Food       2.1288   0.2001  10.64  0.000


S = 12.5559   R-Sq = 27.9%   R-Sq(adj) = 27.7%
```

(a) What are the estimated intercept and slope?

(b) Use the estimated regression model to estimate the average dinner price of all restaurants with a quality rating of 20.

(c) In the estimated regression model, what is the interpretation of the slope?

(d) In the estimated regression model, why doesn't the intercept have a direct interpretation?

7. Refer to the Minitab output from the previous problem, the regression analysis of the Zagat data.

    (a) What is the estimated standard deviation or the error? What is the interpretation of this value?

    (b) According to the estimated regression model, what is the range of typical prices for restaurants with quality ratings of 20?

    (c) According to the estimated regression model, what is the range of typical prices for restaurants with quality ratings of 10?