Hans Trautlein
Due: February 5, 2016

Problem Set 2: Chapter 3
Andrew Bray

Book Problems

1. *Describe the null hypotheses to which the p-values given in Table 3.4 correspond. Explain what conclusions you can draw based on these p-values. Your explanation should be phrased in terms of* `sales`*,* `TV`*,* `radio`*, and* `newspaper`*, rather than in terms of the coefficients of the linear model.*

   TV The null hypothesis here is that spending on television advertising didn't affect sales. The p-value here is very small, so we can reject the null hypothesis at a significant level of $\alpha = 0.05$. This means that television advertising does have an affect on sales.

   Radio The null hypothesis here is that spending on radio advertising didn't affect sales. The p-value here is very small, so we can reject the null hypothesis at a significant level of $\alpha = 0.05$. This means that radio advertising does have an affect on sales.

   Newspaper The null hypothesis here is that spending on newspaper advertising doesn't affect sales. The p-value, at `0.8599` means that we fail to reject the null hypothesis, which in this circumstance implies that newspaper advertising doesn't affect sales.

5. *Consider the fitted values that result from performing linear regression without an intercept. In this setting, the ith fitted value takes the form*

$$\hat{y}_i = x_i \hat{\beta},$$

   where

$$\hat{\beta} = \left(\sum_{i=1}^{n} x_i y_i \right) / \left(\sum_{i'=1}^{n} x_{i'}^2 \right).$$

   Show that we can write

$$\hat{y}_i = \sum_{i'=1}^{n} a_{i'} y_{i'}.$$

   What is $a_{i'}$?

   Note: We interpret this result by saying that the fitted values from linear regression are linear combinations of the response values.

6. *Using (3.4), argue that in the case of simple linear regression, the least squares line always passes through the point (x?, y?).*

Challenge Problem

Use the identities for expected value and variance to derive the bias-variance decomposition of

$$E\left[\left(y - \hat{f}(x)\right)^2\right].$$

$E(c) = c$, where c is constant.
$Var(X) = E(X^2) - [E(X)]^2$