1. Consider the OOP study described at the top of p.281, which was actually a bit different from the description in our book:

\[
\text{mean } Y = \beta_0 + \beta_1 X^{(1)} + \beta_2 X^{(2)} + \beta_3 X^{(1)}X^{(2)} \quad (1)
\]

The results were:

<table>
<thead>
<tr>
<th>coef</th>
<th>betahat</th>
<th>std.err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>4.37</td>
<td>0.23</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.49</td>
<td>0.07</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.56</td>
<td>1.57</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-0.13</td>
<td>-1.34</td>
</tr>
</tbody>
</table>

(a) (10) The last term in (1) is known as the _______________ term.

(b) (20) Find the estimated difference in mean completion time under OOP and using procedure language (former minus the latter), for 1000-line programs.

(c) (15) Find an approximate 95% confidence interval for $\beta_1$, answering with R’s `c()` form.

(d) (15) Find $\hat{\text{Var}}(\hat{\beta}_0)$.

2. (15) In the marbles example, p.147, find $m_{Y:B}(2)$.

3. The code below estimates the regression function $m_{Y:X}(t)$ for scalar $X$, without assuming a linear or other parametric model. The vector parameters $y$, $x$, and the scalar parameter $t$, are self-explanatory. As to the scalar parameter $h$, I’ll simply say that we consider one number $u$ “near” another number $v$ if $|u - v| < h$.

\[
\text{nonparregest} \leftarrow \text{function}(y,x,t,h) \{ \\
\text{dists} \leftarrow \text{abs}(x-t) \\
\text{xnear} \leftarrow \text{which}(\text{dists} < h) \\
\text{mean}(y[\text{xnear}]) \\
\}
\]

(5) Fill blank (a).

(10) Fill blank (b).

(10) Fill blank (c).

Solutions:

1.a interaction

1.b

They also used logarithms, but we’ll ignore that here.