

Name: \_\_\_\_\_

Directions: **Work only on this sheet** (on both sides, if needed). MAKE SURE TO COPY YOUR ANSWERS TO A SEPARATE SHEET FOR SENDING ME AN ELECTRONIC COPY LATER.

**Important note:** Remember that in problems calling for R code, you are allowed to use any built-in R function, e.g. `choose()`, `sum()`, etc.

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

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

The results were:

coef.	betahat	std.err.
$\beta_0$	4.37	0.23
$\beta_1$	0.49	0.07
$\beta_2$	0.56	1.57
$\beta_3$	-0.13	-1.34

- (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  $\widehat{Var}(\hat{\beta}_0)$ .

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  $|u - v| < h$ .

```
nonparregest <- function(y,x,t,h) {  
  dists <- blank (a)  
  xnear <- blank (b)  
  blank (c)  
}
```

- (5) Fill blank (a).
- (10) Fill blank (b).
- (10) Fill blank (c).

**Solutions:**

**1.a** interaction

**1.b**

```
(4.37 + 0.49*1000 + 0.56*1 - 0.13*1000*1) -  
(4.37 + 0.49*1000 + 0.56*0 - 0.13*1000*0)
```

**1.c**

```
c(0.49 - 1.96 * 0.07, 0.49 + 1.96 * 0.07)
```

**1.d**

```
0.23 ^ 2
```

**2.**

```
(0.036*0 + 0.048*1 + 0.006*2) / (0.036 + 0.048 + 0.006)
```

**3.**

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

<sup>1</sup>They also used logarithms, but we'll ignore that here.