11-1

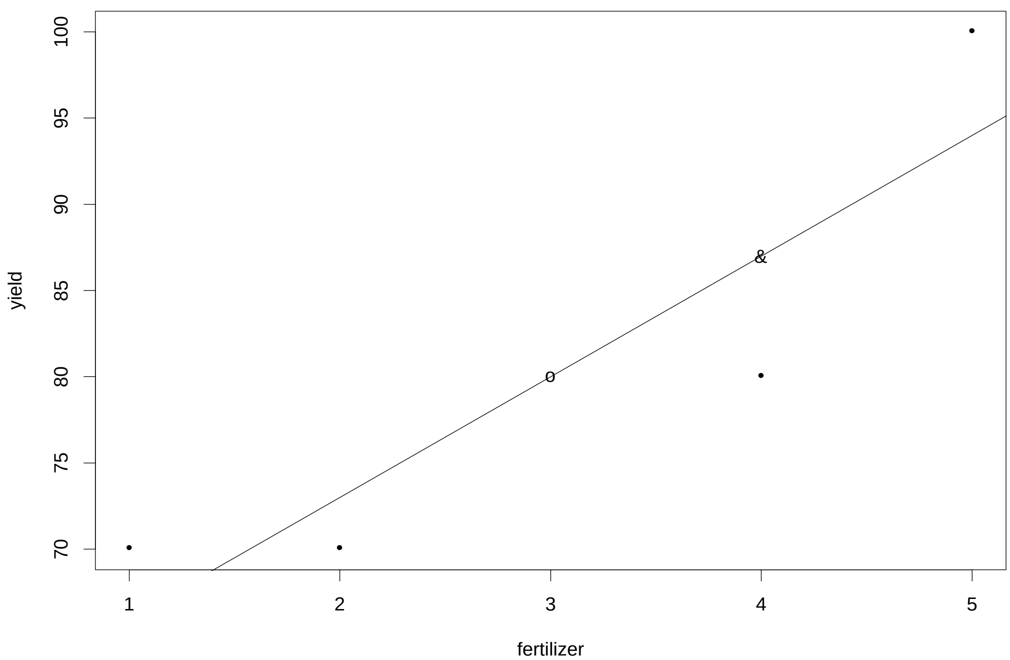
a. For the given data, $\bar{X} = 3$, $\bar{Y} = 80$, $\sum xy = 70$ and $\sum x^2 = 10$, so that

$$b = \frac{\sum xy}{\sum x^2} = \frac{70}{10} = 7$$
$$a = \bar{Y} - b\bar{X} = 80 - 7 \cdot 3 = 59 \ ,$$

and the estimated regression line is

$$\hat{Y} = 59 + 7X \ .$$

b. Here is the graphic. The only points of interest here are the ones marked with a "•" .



c.

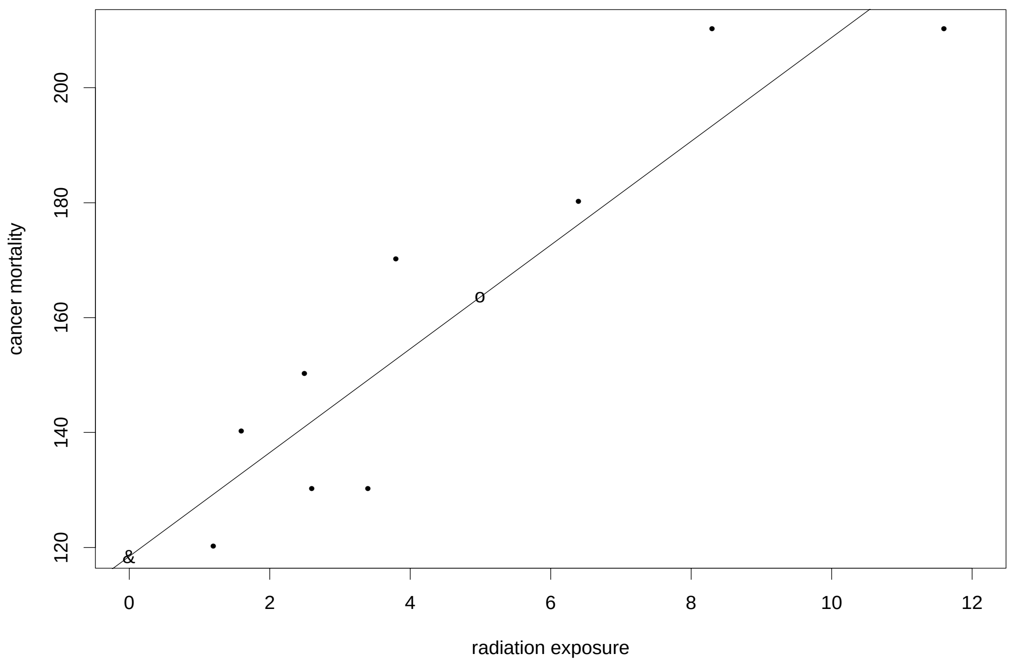
- **i.** $\hat{Y} = 59 + 7 \cdot 3 = 80$; this corresponds to the point marked with an "o" in the previous graph.
- **ii.** $\hat{Y} = 59 + 7 \cdot 4 = 87$; this corresponds to the point marked with an "&" in the previous graph.
- **ii.** This corresponds to the estimated slope of the regression line, $b = 7$.

11-3

a. Here, $b = \sum xy / \sum x^2 = 876/97 \approx 9.031$ and $a = \bar{Y} - b\bar{X} \approx 160 - 9.031 \cdot 4.6 = 118.457$, so that the estimated regression line is $\hat{Y} = 118.457 + 9.031\ X$.

b. For a radiation exposure of 5.0, we estimate the cancer mortality to be $\hat{Y} = 118.457 + 9.031 \cdot 5 = 163.612$. For a radiation exposure of zero, the estimated cancer mortality is the estimated intercept, $a = 118.457$.

c. The answer to **b.** with zero exposure is marked with an "&", the answer to part **b.** with exposure 5.0 with an "o" and the counties with a "•".



d. Since the data arises from an uncontroled observational study, one can not conclude for any causal relationship between radiation exposure and cancer mortality. It can be the case that the observed positive relation between the two variables is due to the presence of some confounding variable.

11-5

a. mean; **b.** normal; **c.** easy; **d.** OLS, curve.

12-13

Since $t = b/SE$ and $SE = .27/t_{.025}$, we get $t = 2.76$. Also the hypotheses is $H_0 : \beta \leq 0$, so $0.0025 < p - value < 0.005$. Note that we used a normal standard table since the sample size is large: $n = 1000$.

Extra Exercise

a) We have $n = 32$, so that $df = 30$ and $t_{.025} = 2.05$.

$$\begin{aligned}
\beta &= b \pm t_{.025}\ SE \\
&= 102.289 \pm 2.05 \cdot 24.23 \\
&= 102.289 \pm 49.43 \\
& \quad 52.86 < \beta < 151.72
\end{aligned}$$

$$\begin{aligned}
\alpha &= a \pm t_{.025}\ SE \\
&= -6553.57 \pm 2.04 \cdot 1661.96 \\
&= -6553.57 \pm 3390.40 \\
& \quad -9943.97 < \alpha < -3163.17
\end{aligned}$$

b) Yes, since the 95% confidence interval for β does not contain zero.

c) The hypotheses to test is $H_0 :\ \beta = 100$, so the t statistic is

$$t = \frac{b - 100}{SE} = \frac{102.289 - 100}{24.23} = 0.094.$$

We want to calculate a two-sided p-value, so, looking at the table, we recognize that $p-value > 2 \cdot 0.25$, that is, the p-value is larger than 50%.

d) Looking at the output tables, we see that $s^2 = 18761.24$.