

Exercise 4

Ex 6 (a) $X_1 =$ "weight of turkey from Virginia"
 $X_2 =$ " " " " Wisconsin"

We assume $X_1 \sim N(\mu_1, \sigma^2)$ and $X_2 \sim N(\mu_2, \sigma^2)$
and that $X_1 \perp X_2$

Then, a 95% CI for $\mu_2 - \mu_1$ is given by

$$(\bar{x}_2 - \bar{x}_1) \pm c. se(\bar{x}_2 - \bar{x}_1)$$

where

$$se(\bar{x}_2 - \bar{x}_1) = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

with

with

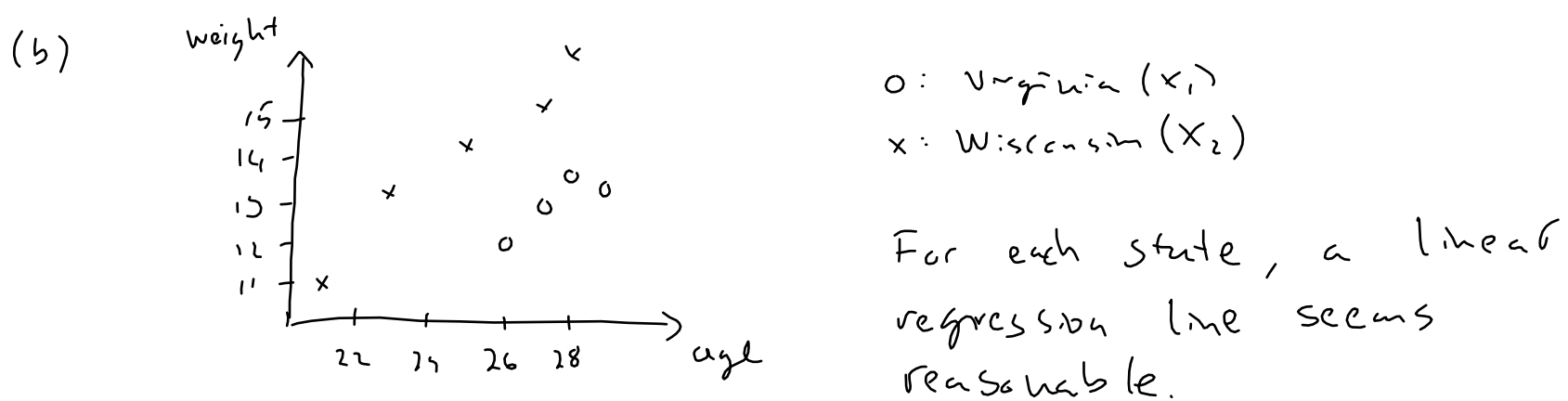
$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

and c is the 97.5 percentile from a $\underbrace{t_{n_1+n_2-2}}$ -distribution

Plugging in the values gives us $s_p \approx 1.17$ and $se(\bar{x}_2 - \bar{x}_1) \approx 0.785$, and from t-table we get $C = 2.365$.

Thus, the interval becomes $(-0.88, 2.83)$. Since the

C1 contains zero, we would not reject

$$H_0: \mu_1 = \mu_2 \quad \text{at sign. level } \alpha = 0.05.$$


(c) β_1 is the average increase in weight per week for a turkey from either state

β_2 is the difference in weight for a turkey from Wisconsin compared to Virginia at a given week.

We assume $E(Y|X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ and that $\epsilon_i \sim N(0, \sigma^2)$.

(d) A 95% CI is given by

$$\hat{\beta}_1 \pm c \cdot se(\hat{\beta}_1)$$

where $c = 2.447$ is obtained as the 97.5 percentile from a t_6 -distribution. The CI then becomes

$$2.094 \pm 2.447 \cdot 0.235 = (1.51, 2.66)$$

and, since the interval does not contain zero, we conclude that weights in Wisconsin and Virginia differ significantly at level $\alpha = 0.05$ when taking age into account.

Exam 2018 - Problem 1

(a) The estimates are
$$\begin{cases} \hat{\beta}_0 = 30.51, \\ \hat{\beta}_1 = -0.08, \\ s = \hat{\sigma} = 2.98. \end{cases}$$

R^2 is the proportion of the variation of Y explained by the regression model.

Since this is a simple linear regression model, we have that $R^2 = r^2$ and, since $\hat{\beta}_1 < 0$, we have that

$$r = -\sqrt{R^2} \approx -0.701$$

(b) The estimate is the same as in (a) because the design is balanced, resulting in uncorrelated covariates.

By including an additional informative covariate R^2 is increased and s is decreased, since the RSS is decreased.

The standard error of $\hat{\beta}_1$ is proportional to s , in the case of uncorrelated covariates, and has thus decreased along with s .