Practice Class Test Model Report

Mitchum Bock

Introduction

Experiments were conducted as part of reseach into "Digitalis", a heart medicine similar to toxins found in plants commonly known as foxglove. 144 domestic male and female adult cats were used in the experiments and they each had their heart weight in grams (`Hwt`) measured. This data, including the gender (`Sex`) of each cat, is analysed in this report.

In particular, this report presents numerical and graphical summaries of the heart weights of the cats and fits a linear model to estimate the difference, on average, between the heart weights of male and female cats.

Exploratory Data Analysis

Summary statistics of the heart weights of the cats are presented in the following table for each sex separately.

Table 1: Summary statistics on heart weight by sex of 144 adult cats.

Sex	n	Mean	St.Dev	Min	Q1	Median	Q3	Max
F	47	9.2	1.4	6.3	8.35	9.1	10.1	13.0
M	97	11.3	2.5	6.5	9.40	11.4	12.8	20.5

This table shows that there were approximately twice as many male cats in the sample (97 compared to 47) and that the summaries of the heart weights of the male cats were consistently greater than the corresponding summaries of the heart weights of female cats. These differences can be easily seen in the following boxplots which summarise the distributions of the heart weights of male and female cats.

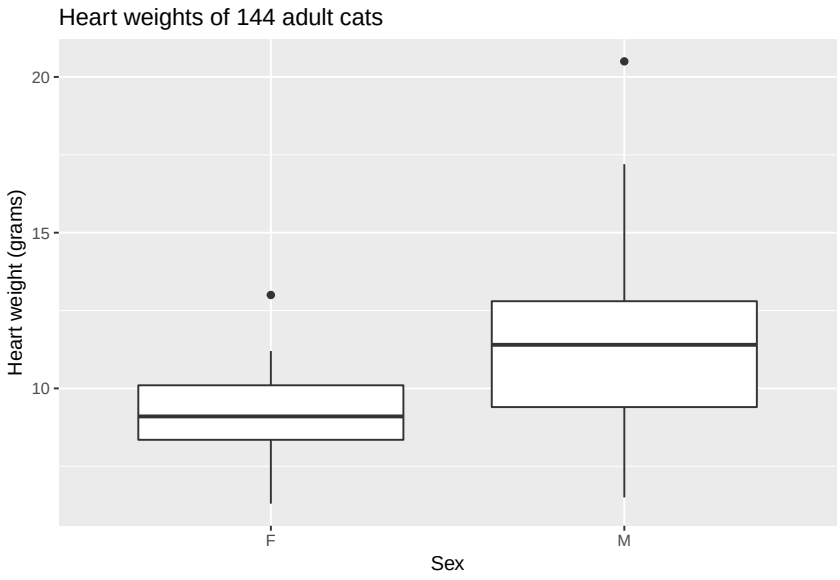

Figure 1: Heart weight by Sex.

The boxplot shows than in addition to the male cats having heavier hearts on average compared to the female cats' hearts, the weights of the male hearts were also more widely distributed. This is also reflected in the standard deviation of the male heart weights being nearly twice the standard deviation of the female cats' heart weights (2.5 grams compared to 1.4 grams, respectively). There are also potentially two outliers (one male and one female) as shown by the two points shown beyond the "whiskers" of the boxplots.

Formal Data Analysis

To begin to analyse the cat heart weights data formally, we fit the following linear model to the data.

$$\widehat{\text{Hwt}}_i = \widehat{\alpha} + \widehat{\beta}_{\text{Male}} \cdot \mathbb{I}_{\text{Male}}(x)$$

where

- $\widehat{\text{Hwt}}_i$ is the expected value of the heart weight of the ith cat in the sample
- the intercept $\widehat{\alpha}$ is the mean heart weight for the baseline category of females;
- $\widehat{\beta}_{\text{Male}}$ is the difference in the mean heart weight of a males relative to the baseline category females; and
- $\mathbb{I}_{\text{Male}}(i)$ is an indicator function such that

$$\mathbb{I}_{\text{Male}}(i) = \begin{cases} 1 & \text{if Sex of } i\text{th observation is Male,} \\ 0 & \text{Otherwise.} \end{cases}$$

When this model is fitted to the data, the following estimates of α (`intercept`) and β_{Male} (`SexM`) are returned:

Table 2: Estimates of the parameters from the fitted linear regression model.

term	estimate
intercept	9.202
SexM	2.121

Hence the model estimates the average heart weight of female cats is 9.202 grams (which agrees with the sample mean reported in Table 1) and that the male cats' heart weights are, on average, 2.121 grams heavier than the female cats' heart weights.

Before we can proceed to use the fitted model (for example to perform statistical inference) we must check the assumptions of the model. These are best considered in light of the residual plots in Figure 2.

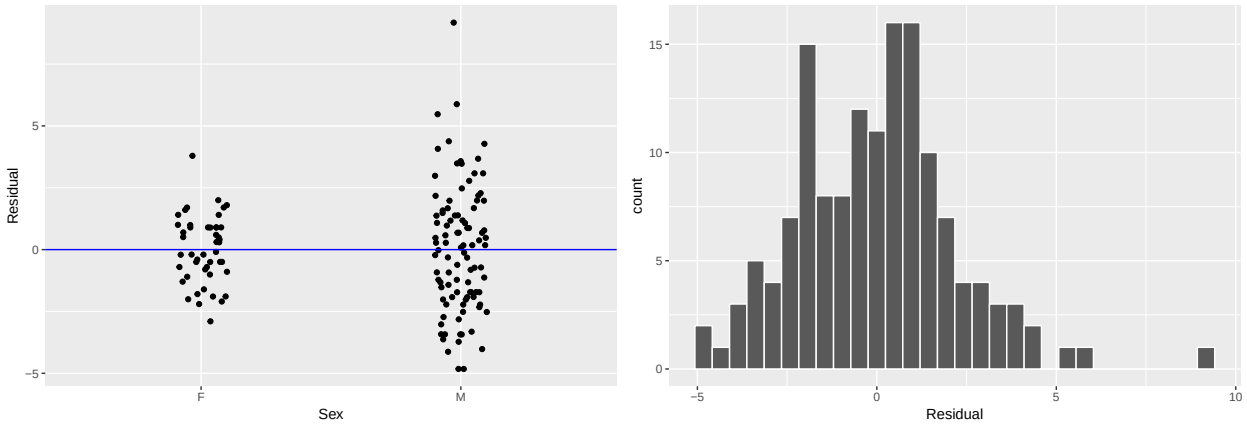

Figure 2: Scatterplots of the residuals by Sex (left) and a histogram of the residuals (right).

The scatterplots show an approximately even spread of the residuals above and below the zero line for each gender, and hence the assumption that the residuals have mean zero appears valid. The assumption of constant variance within the two genders is not supported, however, as the spread of the residuals in the vertical scatter of the male cats is considerably more than that of the females (as was noted above when the standard deviations were considered). The histogram supports the assumption of normally distributed errors in the model, with the exception of a potential outlier.

Conclusions

In summary, we have estimated that, on average, the male cats have hearts which weigh 2.121 grams more than the female cats' hearts. In particular, we estimate the average heart weight of female cats is 9.202 grams and the average heart weight of male cats is 11.3 grams.

In addition to the centers of the distibutions of male and female cats' heart weights being different, we have also observed that the spread of the male heart weights is greater than the spread of the female cats' heart weights. This may pose a problem if the standard linear model was used to further analyse this data, and therefore it is recommended that models which allow for differences in the variances within different groups be used.