Topic 2 Exercise

1. Each female horseshoe crab has a male crab attached to her in her nest. There is a study investigated factors that affect whether the female crab had any other males, called satellites, residing near her.

Data fields:

crab: crab id

sat: number of satellites

y: had or had no satellites

weight: weight of the crab in kilogram

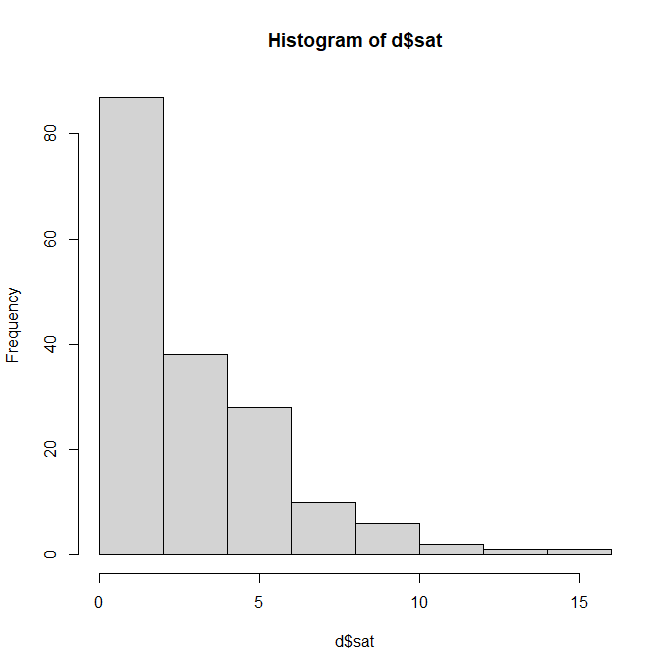
width: shell width in centimeter

color: crab color – 1, medium light; 2, medium; 3, medium dark; 4, dark

spine: spine condition – 1, both good; 2, one broker; 3, both broken

1. Read the text data file “Crabs.dat” into R. Describe the data.

|  |  |
| --- | --- |
| Satellites | Number |
| Y | 111 |
| N | 62 |



1. Identify the factors that affect whether the female crab had any satellites residing near her (Using logistic regression).

* Univariate analysis resulted in significant factors of weight, width, and color. Spine condition is not significant.
* Stepwise selection using BIC resulted in univariate model (~Width) to be the best fit.
* Stepwise selection using AIC resulted in multivariate model (~Width+color) to be the best fit.
* Hosmer-Lemeshow test is not significant, which indicates good fit.

1. Identify the factors that affect whether the number of satellites the female crab had residing near her (Using Poisson regression and negative binomial regression).

* Univariate analysis resulted in significant factors of weight, width, spine, and color.
* Stepwise selection using BIC resulted in univariate model (~Weight) to be the best fit.
* Stepwise selection using AIC resulted in multivariate model (~Weight+color) to be the best fit.
* Over-dispersion test is significant, which indicates negative binomial regression fits better than Poisson regression.