Analysis of the Cereal Dataset

Information on the distribution of various groups of nutrients among 77 commonly available breakfast cereals was obtained from the statistical graphics exposition in 1993. The amount of sugar/serving and shelf position variables were recorded.

We analyzed if there is evidence that the average amount of sugar/serving varies by shelf height. A simple regression of sugars vs. shelf is not appropriate because the shelf should be a categorical variable (or factor) instead of a numerical variable. So first of all, we should convert the shelf variable to a factor. From the formal hypothesis test, we got F-statistic = 6.601 on 2 and 73 DF, p-value=0.002316. So there are strong evidence that the average amount of sugar/serving varies by shelf height.

In figure 1, we found that the highest mean weight of sugar/serving is in the meddle shelf. The confidence interval of middle shelf do **NOT overlap** with others, so there **is** evidence that the mean weight of sugar/serving in middle shelf is different from the other shelves. However, the confidence intervals of the low shelf and high shelf **overlap**, so there **is NO** evidence that the mean weight of sugar/serving are different between low shelf and high shelf. The estimated mean amount of sugar in each shelf and their 95% CI shown in table 1.

shelfCFO	lsmean	SE	df	lower.CL	upper.CL
low	5.11	0.937	73	3.24	6.97
medium	9.62	0.891	73	7.84	11.40
high	6.53	0.681	73	5.17	7.88

Confidence level used: 0.95

Table 1. Estimated mean amount of sugar/serving and 95% confidence interval

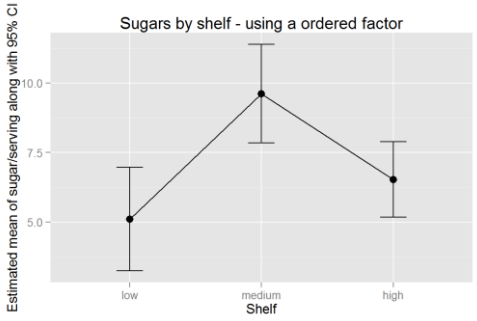

Figure 1. plot of the estimated mean amount of sugar/serving and 95% confidence interval.