Assignment 11, Introduction to Statistics

Oleg Sivokon

<2015-03-27 Fri>

Contents

1 Problems

1.1 Problem 1

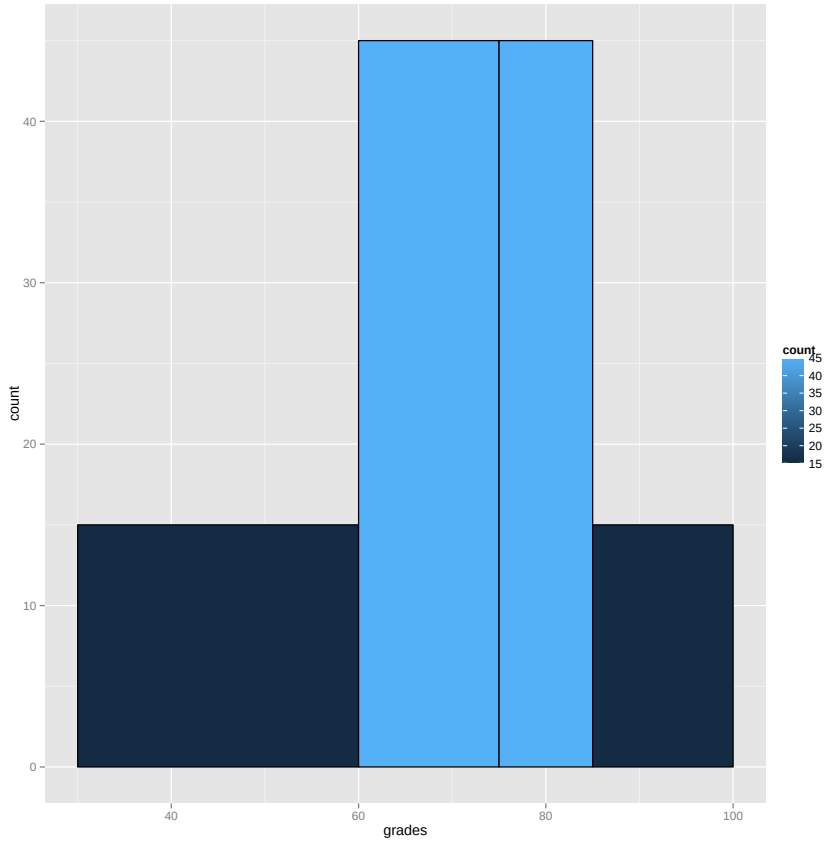
Given the grades in an engineering faculty were as follows:

lower	higher	graded
30	60	15
60	75	45
75	85	45
85	100	15

1. Present the data using a histogram.
2. Calculate mode, median, algebraic average and variance.
3. Calculate the number of students who earned at least 82 points.
4. Given the following data for the preceding year for 80 students where the average grade was 70 and variance was 200, find the average and the variance for two years combined.

1.1.1 Answer 1

```
library(ggplot2)
tbl$avg <- (tbl$lower + tbl$higher) / 2
ggplot(tbl) +
    geom_histogram(
        aes(x = avg, weight = graded, fill = ..count..),
        breaks = unique(append(tbl$lower, tbl$higher)),
        position = "identity", colour = "black") +
            xlab("grades")
```



1.1.2 Answer 2

As easy to see from the diagram, there are two **modes**: one is in the (60, 75] range, an another is in the (75, 85] range.

The **median** is the value at the 60'th studen, which is easy to see from

the table as being the 75 points.

$$
\begin{aligned}
Md &= \frac{\frac{n}{2} - F(x_{m-1})}{f(x_m)} * (L_1 - L_0) + L_0 \\
Md &= \frac{\frac{120}{2} - 60}{45} * (85 - 75) + 75 \\
Md &= \frac{60 - 60}{45} * (85 - 75) + 75 \\
Md &= 75
\end{aligned}
$$

The **average** is given by the formula:

$$
\begin{aligned}
\frac{15 * \frac{30+60}{2} + 45 * \frac{60+75}{2} + 45 * \frac{75+85}{2} + 15 * \frac{85+100}{2}}{15 + 45 + 45 + 15} &= \\
\frac{15 * 90 + 45 * 135 + 45 * 160 + 15 * 185}{2 * 120} &= \\
\frac{17400}{240} &= \\
72.5
\end{aligned}
$$

And the **variance**:

$$
\begin{aligned}
S^2 &= \frac{\sum_1^n (x - \overline{x})^2 * f(x)}{n} \\
S^2 &= \frac{(45 - 72.5)^2 * 15 + (67.5 - 72.5)^2 * 45 + (80 - 72.5)^2 * 45 + (92.5 - 72.5)^2 * 15}{120} \\
S^2 &= \frac{756.25 * 15 + 25 * 45 + 56.25 * 45 + 400 * 15}{120} \\
S^2 &= \frac{21000}{120} \\
S^2 &= 175
\end{aligned}
$$

1.1.3 Answer 3

Using a slightly altered formula for the median, we can calculate the 82'nd percentile. It is easy to see that the 82'nd percentile falls in the third group, viz. [75, 85) interval. Assuming values are uniformly distribute in this

interval, $\frac{7}{10}$ of these are below 82 and $\frac{3}{10}$ are above. In other words, we need to take 0.3 of the 45 students in this category, i.e. 13.5, together with 15 students who earned more points this gives roughly 19 students.

1.1.4 Answer 4

New average is just the weighted average of both averages: $\frac{70*80+72.5*120}{80+120} = 71.5$.

The total variance is calculated using

$$
\begin{aligned}
s_1 &= n(S_1 + \overline{x}_1^2) \\
s_1 &= 120 * (175 + 72.5^2) \\
s_1 &= 651750. \\
s_2 &= n(S_2 + \overline{x}_2^2) \\
s_2 &= 80 * (200 + 70^2) \\
s_2 &= 408000.
\end{aligned}
$$

$$
\begin{aligned}
S^2 &= \frac{s_1 + s_2}{n_1 + n_2} - \overline{x}^2 \\
S^2 &= \frac{651750 + 408000}{200} - 71.5^2 \\
S^2 &= \frac{651750 + 408000}{200} - 5112.25 \\
S^2 &= 186.5
\end{aligned}
$$