PhaseIII

Jiayi Gao

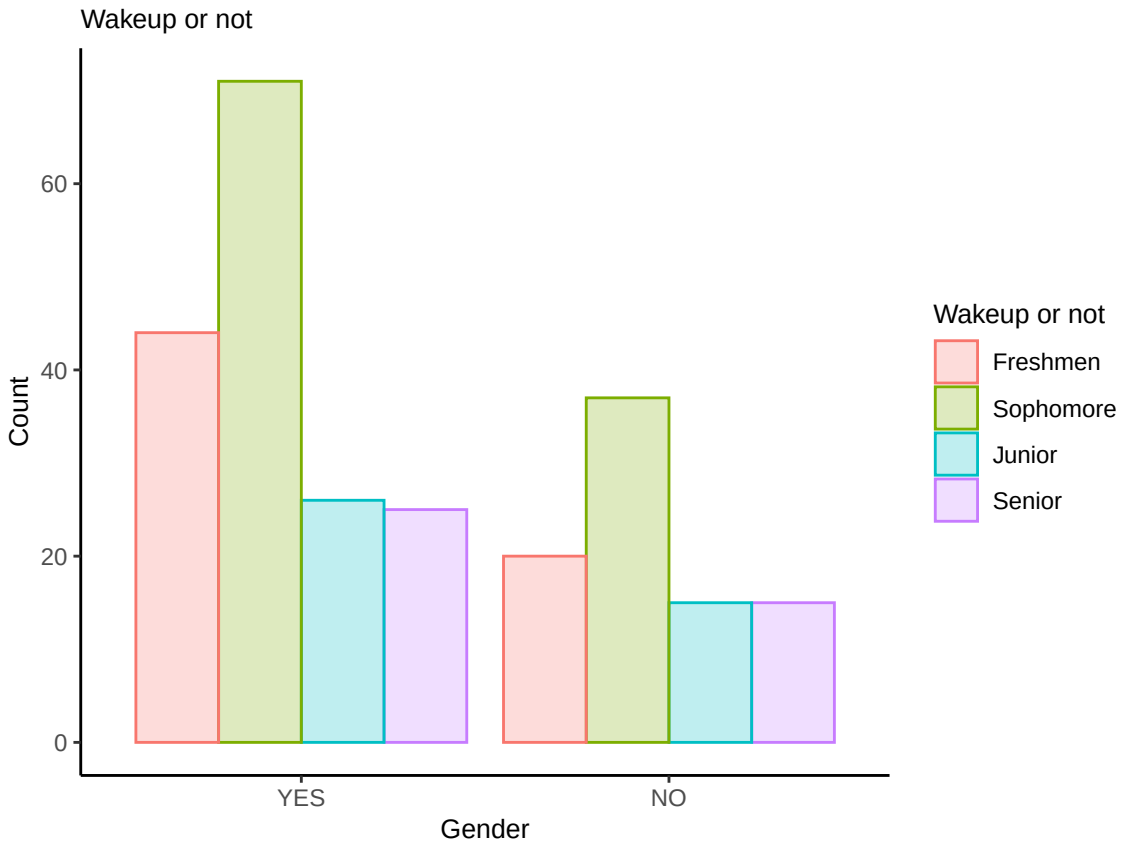
2019/12/19

R Markdown

Thirdly, to test if there exists a significant association between two variables, a chi-square test for association is performed.The null hypothesis for this test is H0: there is not an association between the two categorical variables, and the alternative hypothesis is Ha:there is an association between two categorical variables. Before performing chi-square test, the study has to check if the expected count of each cell is greater than 5. After checking the expected count, the study can put the expected count into a two way table that exactly matches the null hypothesis of no relationship. The remaining task is to calculate the chi-square statistic by using the formula $X^2 = \sum \frac{(Observed - Expected)^2}{Expected}$. Before the study can find p-value, the degree of freedom is needed. By using (row-1)*(column-1), the study can find the degree of freedom. The p-value for the test can be found by using the upper tail of a chi-square distribution with the corresponding degrees of freedom. Finally, if the p-value is less than the significance level α (usually 0.05), the study has fairly strong evidence that there is an association between the two variables.

H_0: School year is associated with Wakeup time
H_A: School year is not associated with Wakeup time

	Freshmen	Sophomore	Junior	Senior
WAKEUP AT NIGHT	13.03	10.51	14.29	19.76
NOT WAKEUP AT NIGHT	17.97	14.49	19.71	27.24




```
  check.all
1      TRUE
```

Year	Wake	obs.counts	exp.counts
Freshmen	YES	44	41.99209
Sophomore	YES	71	70.86166
Junior	YES	26	26.90119
Senior	YES	25	26.24506
Freshmen	NO	20	22.00791
Sophomore	NO	37	37.13834
Junior	NO	15	14.09881
Senior	NO	15	13.75494

Year	Wake	obs.counts	exp.counts	X.sq
Freshmen	YES	44	41.99209	0.0960105
Sophomore	YES	71	70.86166	0.0002701
Junior	YES	26	26.90119	0.0301896
Senior	YES	25	26.24506	0.0590653
Freshmen	NO	20	22.00791	0.1831925
Sophomore	NO	37	37.13834	0.0005153
Junior	NO	15	14.09881	0.0576031
Senior	NO	15	13.75494	0.1126993

	Freshmen	Sophomore	Junior	Senior	TOTAL
WAKEUP AT NIGHT	44	71	26	25	166
NOT WAKEUP AT NIGHT	20	37	15	15	87
TOTAL	64	108	41	40	253

```
    Pearson's Chi-squared test

data:  x
X-squared = 4.2432, df = 3, p-value = 0.2364
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.