Lab 7

Wesley Wessen

11/23/2020

In his second experiment, Mendel differentiated his pea plants by three characteristics, 'form', 'albumen', and 'seed-coat'. Form here could be round (A) or wrinkled (a), albumen yellow (B) or green (b), and seed-coat grey-brown (C) or white (c). Thus, each plant can be categorized according to whether it displays one or the other of each of these characteristics, or according to whether it shows a mix of these characteristics as a hybrid. For example, then, 'ABC' would characterize peas that are yellow and round with grey-brown seed coat, and 'AaBc' would characterize peas having a mix of round yellow with white seed-coat and wrinkled yellow with white seed-coat types.

Evaluate Mendel's claim on the basis of his observed counts, using significance level 0.1 . To do so, first report a table of observed values, expected values, and discrepancies between them. Which categories appear most discrepant? What is the Null distribution for the test statistic here (including its degrees of freedom)? Would you reject the null hypothesis here, or fail to do so?

Found at https://library.si.edu/digital-library/book/versucheberpflan00mend (image is also on canvas)

Hypothesized Ratios

```
## Warning: package 'pander' was built under R version 4.0.3
```

Table 1: Relative ratios under the Mendel's proposal (continued below)

Category	ABC	ABc	AbC	Abc	aBC	aBc	abC	abc	ABCc	AbCc
NullRatio	1	1	1	1	1	1	1	1	2	2

Category	aBCc	abCc	ABbC	ABbc	aBbC	aBbc	AaBC	AaBc	AabC
NullRatio	2	2	2	2	2	2	2	2	2

Category	Aabc	ABbCc	aBbCc	AaBCc	AabCc	AaBbC	AaBbc	AbBbCc
NullRatio	2	4	4	4	4	4	4	8

Observed, Expected, and Discrepancies

Table 4: Observed vs. Expected values for Experiment 1 (continued below)

	ABC	ABc	AbC	Abc	aBC	aBc
NullProportion	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156
Expected	9.98	9.98	9.98	9.98	9.98	9.98
Observed	8	14	9	11	8	10
Discrepancy	0.394	1.62	0.0971	0.103	0.394	2.45e-05

	abC	abc	ABCc	AbCc	aBCc	abCc
NullProportion	0.0156	0.0156	0.0312	0.0312	0.0312	0.0312
Expected	9.98	9.98	20	20	20	20
Observed	10	7	22	17	25	20
Discrepancy	2.45e-05	0.892	0.207	0.441	1.27	4.89e-05

	ABbC	ABbc	aBbC	aBbc	AaBC	AaBc
NullProportion	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312
Expected	20	20	20	20	20	20
Observed	15	18	19	24	14	18
Discrepancy	1.24	0.194	0.047	0.814	1.78	0.194

	AabC	Aabc	ABbCc	aBbCc	AaBCc	AabCc
NullProportion	0.0312	0.0312	0.0625	0.0625	0.0625	0.0625
Expected	20	20	39.9	39.9	39.9	39.9
Observed	20	16	45	36	38	40
Discrepancy	4.89e-05	0.789	0.642	0.388	0.094	9.78e-05

	AaBbC	AaBbc	AbBbCc
NullProportion	0.0625	0.0625	0.125
Expected	39.9	39.9	79.9
Observed	49	48	78
Discrepancy	2.06	1.63	0.044

Test Statistic and p-value

```
## [1] 0.9511498
```

Discussion

The categories with the most discrepancies are AaBbC, AaBC, AaBbc, ABc, aBCc, and ABbC. These all have discrepancies of over 1, meaning the numbers observed for these categories deviate significantly from the expected values. Categories with discrepancies of over 0.5 include abc, aBbc, Aabc, and ABbCc, these values also have abnormally large observed numbers compared to the expected. With 27 categories, there

are 26 degrees of freedom. The chi-squared test results in a p-value of 0.9511498, which is very large, so there is a lack of evidence against the null hypothesis, and that the discrepancies are a result of random chance.

Appendix

```
knitr::opts_chunk$set(echo = FALSE)
library(pander)
Category <- c("ABC","ABc","AbC","Abc",
              "aBC","aBc","abC","abc",
              "ABCc","AbCc","aBCc","abCc",
              "ABbC","ABbc","aBbC","aBbc",
              "AaBC","AaBc","AabC","Aabc",
              "ABbCc","aBbCc",
              "AaBCc","AabCc",
              "AaBbC","AaBbc",
              "AbBbCc")
NullRatio <- c(1, 1, 1, 1,
               1, 1, 1, 1,
               2, 2, 2, 2,
               2, 2, 2, 2,
               2, 2, 2, 2,
               4, 4,
               4, 4,
               4, 4,
               8)
ATable <- rbind(Category,NullRatio)
panderOptions('table.continues', '')
pander(ATable, caption="Relative ratios under the Mendel's proposal")
NullProportion <- NullRatio/sum(NullRatio) # converting ratios to proportions
Observed <- c(8,14,9,11,8,10,10,7,22,17,25,20,15,18,19,24,14,18,20,16,45,36,38,40,49,48,78) # from Mend
Expected <-sum(Observed)*NullProportion # converting proportions to Expected counts
Discrepancy <- (Expected-Observed)^2/Expected

Table <- rbind(NullProportion,Expected,Observed,Discrepancy)
pander(Table,
      col.names=Category,
      digits=3,
      caption="Observed vs. Expected values for Experiment 1")

TestStat <- sum(Discrepancy) # chi-squared test statistic
p_value <- 1-pchisq(TestStat,26) # don't forget the degrees of freedom!
p_value
```