StatCrunch Lab 12 for Statistics 301

Topics: goodness of fit, test of independence/homogeneity

Goodness of Fit Test.
Given dataset of 556 *observed* frequencies of various types of peas, test whether or not proportion of round–yellow, wrinkled–yellow, round–green and wrinkled–green peas occurs with *equal* frequency or not at 5%.

type	round–yellow	wrinkled–yellow	round–green	wrinkled–green
frequency	315	101	108	32

Blank data table. Relabel var1 observed pea, var2 expected pea proportion. Type data into these two columns. Data, Data expression, Expression: expected pea proportion*1000, New column name: expected pea, Compute. Data, Save data, 12.1.3 pea type distribution.

1. *Statement.*
   $H_0 : p_1 = \frac{1}{4}, p_2 = \frac{1}{4}, p_3 = \frac{1}{4}, p_4 = \frac{1}{4}$
   versus H_1 : observed pea type distribution different from null

2. *Test.*
$$\text{p--value} = P(\chi^2 \geq 322) = 0.00$$

   Stat, Goodness-of-fit, Chi-square test, Observed: observed pea, Expected: expected pea, Calculate.
   Level of significance $\alpha = 0.05$

3. *Conclusion.*
   Since p–value $= 0.00 < \alpha = 0.05$, reject null: H_0 : observed different from expected ratio.

Test of Independence.
Random sample of college attendance by fathers and their oldest sons in a midwestern city recorded in table below. Test[1] whether or not a son attends college is *dependent* on whether or not father attends college at $\alpha = 0.01$.

observed, O_i	son attended college	son did not attend college	
father attended college	18	12	30
father did not attend college	22	33	55
	40	45	85

Blank data table. Relabel var1 father, var2 son attends, var3 son does not attend. Type data into these three columns. Data, Save data, 12.2.1 father and son data.

1. *Statement.*
   H_0 : son attends independent of father attending
   versus H_1 : son attends dependent on father attending

2. *Test.*
$$\text{p–value} = P(\chi^2 \geq 3.15) = 0.08$$

   Stat, Tables, Contingency, with summary, Select columns for table: son attends, son does not attend, Row labels in: father, Next, choose Expected number, Chi-Square, Calculate.
   Level of significance $\alpha = 0.01$

3. *Conclusion.*
   Since p–value $= 0.08 > \alpha = 0.01$, do not reject null H_0 : independence.

A test of homogeneity is conducted in exactly the same was as a test of independence. If both row and column marginal totals are random, the test is called a test of independence; if one of the marginal totals is fixed and the other random, then the test is called a test of homogeneity.

---

[1]No matter how this question is worded, null hypothesis for test is *always* independent and alternative hypothesis is *always* dependent.