Raw data for Section 7.2: Two categorical variables

Var. 1	Var. 2
A	Independent
A	Dem
A	Dem
B	Rep
B	Dem
B	Rep
C	Ind
C	Rep
C	Dem
C	Ind.

Note: If both are binary, you can do 2-proportion procedures instead of chi-square

Raw data for Chapter 8

Var. 1	Var. 2 (Change to quantitative)
A	15
A	18
A	17
B	14
B	11
B	13
C	16
C	17
C	19
C	17

Note: If the categorical variable is binary, you can do 2-sample t procedures instead.

For this setting, need a new statistic. Some preliminary quantities

$$SSG \left(\text{sum-of-squares group}\right) = \sum n_i \left(\bar{x}_i - \bar{x}\right)^2$$

$\bar{x}$ = overall mean

$$= \left(15 + 18 + 17 + 14 + 11 + 13 + 16 + 17 + 19 + 17\right) \cdot \frac{1}{10}$$

$$= 15.7$$

$$\bar{x}_A = (15+18+17)\cdot\frac{1}{3} = 16.67, \quad n_A = 3$$

$$\bar{x}_B = (14+11+13)\cdot\frac{1}{3} = 12.67, \quad n_A = 3$$

$$\bar{x}_C = (16+17+19+17)\cdot\frac{1}{4} = 17.25, \quad n_C = 4$$

$$SSG = \sum n_i(\bar{x}_i - \bar{x})^2$$

$$= n_A(\bar{x}_A - \bar{x})^2 + n_B(\bar{x}_B - \bar{x})^2 + n_C(\bar{x}_C - \bar{x})^2$$

$$= 3(16.67 - 15.7)^2 + 3(12.67 - 15.7)^2 + 4(17.25 - 15.7)^2$$

$$= 39.9754$$

$$SSE \text{ (Sum of squared error)} = \sum(x - \bar{x}_i)^2$$

In Group A:

$$(15 - 16.67)^2 = 2.7889$$
$$(18 - 16.67)^2 = 1.7689$$
$$(17 - 16.67)^2 = 0.1089$$

In Group B:

$$(14 - 12.67)^2 = 1.7689$$
$$(11 - 12.67)^2 = 2.7889$$
$$(13 - 12.67)^2 =$$

In Group C:

$$(16 - 17.25)^2 =$$
$$(17 - 17.25)^2 =$$
$$(19 - 17.25)^2 =$$
$$(17 - 17.25)^2 =$$

total 14.083 = SSE

SST (sum of squares total) $= \sum (x - \bar{x})^2$

$= 54.1$

$SSG + SSE = SST$

	df	SS	MS	F
Group	2	SSG	$SSG/2 = MSG$	$\frac{MSG}{MSE}$
Error	7	SSE	$SSE/7$ MSE	
Total	9	SST		