

1. (a) Not stopped Bribe requested Stopped/Given warning

Upper class 14 6 7 27

Lower class 7 7 1 15
21 13 8 42

$$27 \times 21 / 42 = 13.5 \quad 27 \times 13 / 42 = 8.357 \quad 27 \times 8 / 42 = 5.142 \\ 15 \times 21 / 42 = 7.5 \quad 15 \times 13 / 42 = 4.642 \quad 15 \times 8 / 42 = 2.857$$

$$(14 - 13.5)^2 / 13.5 = 0.0185 \quad (6 - 8.357)^2 / 8.357 = 0.6647 \quad (7 - 5.142)^2 / 5.142 = 0.6713 \\ (7 - 7.5)^2 / 7.5 = 0.0333 \quad (7 - 4.642)^2 / 4.642 = 1.1977 \quad (1 - 2.857)^2 / 2.857 = 1.207$$

$$\chi^2 = 3.7925$$

(b) `pchisq(3.7925, df=2, lower.tail=FALSE)`
[1] 0.1501306

The p-value returned is not significant when $\alpha=0.1$ as it falls below the critical value.

(c) Not stopped Bribe requested Stopped/Given warning

Upper class 0.3220306 -1.641957 1.523026

Lower class -0.3220306 1.641957 -1.523026

The values of the standardised residuals are low, demonstrating that the variables are independent.

2. (a) H_0 : There is no relationship between the policy of reserving the position of GM for women and the number of new or repaired drinking water facilities.

H_a : The policy of reserving GM for women will be associated with an increase/decrease in the number of new or repaired drinking water facilities.

(b) `lm(formula = water ~ reserved, data = data)`

Residuals:

Min 1Q Median 3Q Max

-23.991 -14.738 -7.865 2.262 316.009

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 14.738 2.286 6.446 4.22e-10 ***

reserved 9.252 3.948 2.344 0.0197 *

Residual standard error: 33.45 on 320 degrees of freedom

Multiple R-squared: 0.01688, Adjusted R-squared: 0.0138

F-statistic: 5.493 on 1 and 320 DF, p-value: 0.0197

(c) A p-value of 0.0197 indicates a statistically significant relationship between reservation policy and the number of new/repaired drinking water facilities, allowing us to reject H_0 . The difference between 0 and 1 in our input variable resulted in an estimated increase of 9.252 in our output variable, providing strong evidence in support of H_a . However, the low r-squared value (0.0138) suggests that our model doesn't explain the bulk of the variability in the data