Dr Isabella Gollini

Probability and Statistics

Solutions 6

1. (i) If p denotes the proportion of individuals in the population who are smokers, we test the null hypothesis $H_0 : p = 1/2$ against the alternative $H_1 : p \neq 1/2$. The sample proportion is $\hat{p} = 230/400 = 0.575$.

The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.575 - 0.5}{\sqrt{(0.5)(0.5)/400}} = (40)(0.075) = 3.00.$$

The p-value for the corresponding two-tail test is

$$p = 2(1 - \Phi(3.00)) = 2(1 - 0.99865) = 0.0027.$$

This p-value is significant at the 1% level. There is very strong evidence to reject the claim that the proportion of smokers in the population is 1/2.

(ii) The estimated population proportion is given by $\hat{p} = 0.575$. A 95% confidence interval for the population proportion is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}\,, \qquad \text{i.e.,} \qquad 0.575 \pm 1.96\sqrt{\frac{(0.575)(0.425)}{400}}$$

This gives 0.575 ± 0.04845 so that, correct to 2 decimal places, the 95% confidence interval is $(0.53, 0.62)$.

Using `R`

```
prop.test(230, 400,
          correct = FALSE)


##
##  1-sample proportions test without continuity correction
##
## data:  230 out of 400, null probability 0.5
## X-squared = 9, df = 1, p-value = 0.0027
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.5260675 0.6225056
## sample estimates:
##     p
## 0.575
```

2. If p_1 denotes the underlying proportion of individuals who die under treatment A and p_2 denotes the underlying proportion of individuals who die under treatment B then we test the null hypothesis $H_0 : p_1 = p_2$ against the alternative $H_1 : p_1 \neq p_2$. We have $\hat{p}_1 = 15/150 = 0.10$ and $\hat{p}_2 = 45/300 = 0.15$. Under H_0 the pooled estimate of the proportion who die is

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{(150)(0.1) + (300)(0.15)}{150 + 300} = \frac{60}{450} = \frac{2}{15} = 0.1333 \ .$$

The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\,(1/n_1 + 1/n_2)}} = \frac{0.1 - 0.15}{\sqrt{(0.1333)(0.8667)\,(1/150 + 1/300)}} = -1.47 \ .$$

The p-value for the corresponding two-tail test is

$$p = 2(1 - \Phi(1.47)) = 2(1 - 0.9292) = 0.1416.$$

This p-value is not significant at the 5% level. There is no strong evidence that that there is any difference between the underlying proportions of patients who die under the two treatments.

Using `R`,

```
prop.test(c(15, 45), c(150, 300),
          correct = FALSE)


##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(15, 45) out of c(150, 300)
## X-squared = 2.1635, df = 1, p-value = 0.1413
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.11274946  0.01274946
## sample estimates:
## prop 1 prop 2
##   0.10   0.15
```