STAT1301 Assignment 4

Caleb Yates s49886351

October 14, 2025

1 Question 4

1.1 Part a)

Let X be the random variable for the number of people received the direct mail strategy and compelted screening. Let Y be the random variable for the number of people who received the education only outreach and completed screening.

The notation p_{X} and p_{Y} represent the population proportion for X and Y respectively.

The null hypothesis is that both population proportions are equal: $\mathrm{H}_0 : p_{\mathrm{X}} = p_{\mathrm{Y}} = p$. The alternative hypothesis is therefore: $\mathrm{H}_1 : p_{\mathrm{X}} > p_{\mathrm{Y}}$.

1.2 Part b)

$\mathrm{X} \sim \mathrm{Bin}(n_{\mathrm{X}}, p_{\mathrm{X}})$ where $n_{\mathrm{X}} = 1415$ so $\mathrm{X} \sim \mathrm{Bin}(1415, p_{\mathrm{X}})$. It is (implicitely) assumed that samples X_i from X follow the distribution of X, hence:

$$\mathrm{X}_i \sim \mathrm{Bin}(n_{\mathrm{X}}, p_{\mathrm{X}})$$

Since $n_{\mathrm{X}} p_{\mathrm{X}} = 505 \gg 5$ and $n_{\mathrm{X}}(1 - p_{\mathrm{X}}) = 910 \gg 5$, the conditions for the Central Limit Theorem (CLT) to apply are met, as well as a suitably large n_{X}. Therefore:

$$\mathrm{X}_i \overset{\text{approx}}{\sim} N(n_{\mathrm{X}} p_{\mathrm{X}}, n_{\mathrm{X}} p_{\mathrm{X}}(1 - p_{\mathrm{X}}))$$

$$\bar{\mathrm{X}} \overset{\text{approx}}{\sim} N(p_{\mathrm{X}}, \frac{p_{\mathrm{X}}(1 - p_{\mathrm{X}})}{n_{\mathrm{X}}})$$

Under H_0:

$$\bar{\mathrm{X}} \sim N(p, \frac{p(1-p)}{n_{\mathrm{X}}})$$

$\mathrm{Y} \sim \mathrm{Bin}(n_{\mathrm{Y}}, p_{\mathrm{Y}})$ where $n_{\mathrm{Y}} = 1408$ so $\mathrm{Y} \sim \mathrm{Bin}(1408, p_{\mathrm{Y}})$. It is (implicitely) assumed that samples Y_i from Y follow the distribution of Y, hence:

$$\mathrm{Y}_i \sim \mathrm{Bin}(n_{\mathrm{Y}}, p_{\mathrm{Y}})$$

Since $n_{\mathrm{Y}} p_{\mathrm{Y}} = 264 \gg 5$ and $n_{\mathrm{Y}}(1 - p_{\mathrm{Y}}) = 1144 \gg 5$, the conditions for the Central Limit Theorem (CLT) to apply are met, as well as a suitably large n_{Y}. Therefore:

$$\mathrm{Y}_i \overset{\text{approx}}{\sim} N(n_{\mathrm{Y}} p_{\mathrm{Y}}, n_{\mathrm{Y}} p_{\mathrm{Y}}(1 - p_{\mathrm{Y}}))$$

$$\bar{\mathrm{Y}} \overset{\text{approx}}{\sim} N(p_{\mathrm{Y}}, \frac{p_{\mathrm{Y}}(1 - p_{\mathrm{Y}})}{n_{\mathrm{Y}}})$$

Under H_0:

$$\bar{\mathrm{Y}} \sim N(p, \frac{p(1-p)}{n_{\mathrm{Y}}})$$