# Research Project

## Computer Lab Exercise 2.2.

---

## Question 2: Hypothesis Test

The t-test statistic for the two groups (when x=0 and x=1) can be defined as follows:

$$t - \text{stat} = \frac{\bar{y}_{x=0} - \bar{y}_{x=1} - (\mu_1 - \mu_0)}{\text{SE}\left(\bar{y}_{x=0} - \bar{y}_{x=1}\right)}$$

Where $\bar{y}$ is the mean for each group, $\mu$ the hypotheses and $\text{SE}(\bar{y}_{x=0} - \bar{y}_{x=1})$ is defined as:

$$\text{SE}(\bar{y}_{x=0} - \bar{y}_{x=1}) = \sqrt{\frac{\sigma^2_{x=0}}{n_{x=0}} + \frac{\sigma^2_{x=1}}{n_{x=1}}}$$

The hypotheses are:

$H_0 : \mu_0 = \mu_1$
$H_1 : \mu_0 \neq \mu_1$

Inside StataSE, following commands can be run to get the desired statistics:

```
sum y if x==0
sum y if x==1
```

This outputs the following results:

$\bar{y}_{x=0} = 0.3038462$
$\bar{y}_{x=1} = 0.38125$

$\sigma^2_{x=0} = (0.4603599)^2$
$\sigma^2_{x=1} = (0.4862005)^2$

$n_{x=0} = 520$
$n_{x=1} = 480$

If the null-hypothesis $H_0$ is assumed, then:
$\mu_0 = \mu_1 \Longrightarrow \mu_1 - \mu_0 = 0$

Therefore:

$$t - \text{stat} = \frac{(0.3038462) - (0.38125)}{\sqrt{\frac{(0.4603599)^2}{520} + \frac{(0.4862005)^2}{480}}}$$

For computing purposes, all variables will be redefined with appropriate names (and rounded to 3 decimal places) and the t-stat defined as a function:

In[1]:=
```
y0 = 0.304
```

In[2]:=
```
y1 = 0.381
```

In[3]:=
```
s0 = 0.460
```

In[4]:=
```
s1 = 0.486
```

In[5]:=
```
n0 = 520
```

In[6]:=
```
n1 = 480
```

In[7]:=
```
t[y0_, y1_, s0_, s1_, n0_, n1_] = (y0 - y1) / Sqrt[s0^2/n0 + s1^2/n1]
```

In[8]:=
```
tstat = t[y0, y1, s0, s1, n0, n1]
```

Out[8]=
```
-2.5681
```

Therefore, the t-statistic is equal to -2.5681

The significance level used is $\alpha = 5\%$ therefore $Z_{\frac{\alpha}{2}} = 1.96$

Because $|\text{t-stat}| > 1.96 \Longrightarrow$ We reject $H_0$.