Hair color

Data

```
hair_color <- read.table('hair_color.txt', header = TRUE, sep = ',', row.names = 1)
attach(hair_color)
head(hair_color)
```

```
##        fair red medium dark jetblack
## male    592 119    849  504       36
## female  544  97    677  451       14
```

```
head(is.na.data.frame(hair_color))
```

```
##         fair   red medium  dark jetblack
## male   FALSE FALSE  FALSE FALSE    FALSE
## female FALSE FALSE  FALSE FALSE    FALSE
```

Hypothesis

We have two independent populations grouped by gender. We would like to check whether gender plays any role in hair color of individuals. We set the hypothesis as follow:

In Null hypothesis we assume that there is no effect whatsoever H0 : Gender has no effect on hair color. There is no statistical difference in two gender groups. Ha : Gender has effect on hair color. There is a statistical difference in two gender groups.

Statistical tests

Since we have categorial data we use Chi-squared. Alternately, we could have used Fisher's test but the datase is too small.

Chi-squared test

```
tests.chi <- chisq.test(hair_color)
tests.chi
```

```
##
##  Pearson's Chi-squared test
##
## data:  hair_color
## X-squared = 10.467, df = 4, p-value = 0.03325
```