Gaze behaviour analysis

Zilla Garama

September 2019

Contents

Introduction

This document presents inferential statistical analyses of gaze behaviour as reported in the paper:

Creating windows to the soul: How eye gaze behaviour can make virtual characters more believable, more socially present and have mental states attributed to them

The OSF form belonging to this paper can be found here: https://osf.io/8u2kf

Authored by Zilla Garama, Ding Ding, and Willem-Paul Brinkman.

The gaze behaviour that has been gathered is the following:

- phase 1:
  - the total duration that the participant looks at each virtual character in this phase
  - how often the participant looks at each virtual character in this phase
- phase 2:
  - the total duration that the participant looks at each virtual character in this phase
  - how often the participant looks at each virtual character in this phase
  - how often the participant follows the gaze of each virtual character to objects in the scene during this phase

Abbreviations:

- virtual character (VC)
- Control virtual character (CVC)
- Gaze aware virtual character (GAVC)

Libraries used:

```
library(ggplot2)        # plotting & data
library(psych)          # reliability function
library(pastecs)        # plotting & data
library(lsr)            # effect size
```

Read in the gaze behaviour data:

```
GB_data <- read.csv("GB_data.csv", header = TRUE, stringsAsFactors = FALSE,
                    fileEncoding="UTF-8-BOM")
GB_data<-GB_data[-c(25:28),]
```

Data preparation

The data is originally saved in the format minutes:seconds. However, the data shows that the minutes are always 00. So, I stripped the minute part off to make it easier to work with the data. At the same time the data is saved in such a way that makes R recognize the data as continous instead of discrete.

```
GB_data$GB_3_phase1_total_time_CVC <- as.numeric(gsub("00:", "",
                                                      GB_data$GB_3_phase1_total_time_CVC))
GB_data$GB_4_phase1_total_time_GAVC <- as.numeric(gsub("00:", "",
                                                       GB_data$GB_4_phase1_total_time_GAVC))
GB_data$GB_7_phase2_total_time_CVC <- as.numeric(gsub("00:", "",
                                                      GB_data$GB_7_phase2_total_time_CVC))
GB_data$GB_8_phase2_total_time_GAVC <- as.numeric(gsub("00:", "",
                                                       GB_data$GB_8_phase2_total_time_GAVC))

print(GB_data)
```

```
##    Participant_number GB_1_CVC GB_2_GAVC GB_3_phase1_total_time_CVC
## 1                 180    UMA_4     UMA_5                          28
## 2                 122   UMA_5S    UMA_4S                          13
## 3                 198    UMA_4     UMA_5                          33
## 4                 182    UMA_5     UMA_4                          27
## 5                 132   UMA_4S    UMA_5S                          40
## 6                 107    UMA_2     UMA_3                          36
## 7                 184    UMA_3     UMA_2                          28
## 8                 137   UMA_4S    UMA_5S                          38
## 9                 103    UMA_3     UMA_2                          21
## 10                136    UMA_3     UMA_2                          33
## 11                128    UMA_4     UMA_5                          35
## 12                185   UMA_3S    UMA_2S                          27
## 13                163   UMA_2S    UMA_3S                          34
## 14                112    UMA_3     UMA_2                          28
## 15                149   UMA_5S    UMA_4S                          43
## 16                173    UMA_3     UMA_2                          47
## 17                195    UMA_2     UMA_3                          37
## 18                115    UMA_4     UMA_5                          14
## 19                169   UMA_4S    UMA_5S                          56
## 20                120   UMA_5S    UMA_4S                          16
```

```
## 21                146   UMA_5S    UMA_4S                          43
## 22                178   UMA_2S    UMA_3S                          29
## 23                119   UMA_3S    UMA_2S                          26
## 24                157   UMA_2S    UMA_3S                          25
##    GB_4_phase1_total_time_GAVC GB_5_phase1_number_of_times_CVC
## 1                           16                               14
## 2                           29                               13
## 3                           38                               25
## 4                           54                               33
## 5                           29                               34
## 6                           36                               37
## 7                           18                               21
## 8                           39                               32
## 9                           29                               24
## 10                          42                               35
## 11                          28                               24
## 12                          27                               23
## 13                          33                               33
## 14                          27                               34
## 15                          40                               46
## 16                          46                               39
## 17                          49                               34
## 18                          27                               24
## 19                          52                               39
## 20                          14                               27
## 21                          37                               32
## 22                          27                               56
## 23                           6                               57
## 24                          27                               35
##    GB_6_phase1_number_of_times_GAVC GB_7_phase2_total_time_CVC
## 1                                14                         10
## 2                                12                          3
## 3                                25                         22
## 4                                39                         19
## 5                                28                         30
## 6                                34                         32
## 7                                17                         31
## 8                                27                         18
## 9                                20                          1
## 10                               32                         26
## 11                               29                         11
## 12                               27                         33
## 13                               34                         23
## 14                               34                         20
## 15                               41                         26
## 16                               37                         25
## 17                               40                         24
## 18                               28                         18
## 19                               38                         17
## 20                               22                         18
## 21                               33                         38
## 22                               46                         21
## 23                               26                          3
## 24                               41                         12
```

```
##    GB_8_phase2_total_time_GAVC GB_9_phase2_number_of_times_CVC
## 1                           11                               10
## 2                            2                               12
## 3                           21                               10
## 4                           16                               23
## 5                           14                               19
## 6                           27                               31
## 7                           13                               18
## 8                           20                               16
## 9                            1                                6
## 10                          30                               32
## 11                          12                               13
## 12                          15                               17
## 13                          19                               28
## 14                          19                               25
## 15                          18                               22
## 16                          23                               23
## 17                          29                               22
## 18                          16                               16
## 19                          23                               23
## 20                          18                               23
## 21                          28                               22
## 22                          19                               38
## 23                           1                               14
## 24                          10                               30
##    GB_10_phase2_number_of_times_GAVC
## 1                                 11
## 2                                  5
## 3                                 15
## 4                                 13
## 5                                 15
## 6                                 22
## 7                                 13
## 8                                 15
## 9                                  3
## 10                                31
## 11                                21
## 12                                19
## 13                                15
## 14                                23
## 15                                21
## 16                                22
## 17                                29
## 18                                15
## 19                                26
## 20                                20
## 21                                30
## 22                                37
## 23                                 8
## 24                                16
##    GB_11_phase2_number_of_times_objects_CVC
## 1                                         5
## 2                                        44
## 3                                         3
```

```
## 4                                          5
## 5                                          3
## 6                                          6
## 7                                          1
## 8                                          2
## 9                                          4
## 10                                         3
## 11                                         4
## 12                                         4
## 13                                         1
## 14                                         3
## 15                                        22
## 16                                         5
## 17                                         3
## 18                                         2
## 19                                         3
## 20                                         3
## 21                                         0
## 22                                         2
## 23                                         3
## 24                                         2
##    GB_12_phase2_number_of_times_objects_GAVC
## 1                                          4
## 2                                          4
## 3                                          6
## 4                                          2
## 5                                          4
## 6                                          4
## 7                                          2
## 8                                          4
## 9                                          5
## 10                                         3
## 11                                         5
## 12                                         3
## 13                                         4
## 14                                         2
## 15                                         2
## 16                                         6
## 17                                         2
## 18                                         5
## 19                                         5
## 20                                         3
## 21                                         4
## 22                                         3
## 23                                         3
## 24                                         1
```

Next, two tables are created, one for the CVC and one for the GAVC. This is to make handling and viewing the data easier.

The data table of the CVC:

```
GB_CVC <- data.frame(Participant_number = GB_data$Participant_number,
                     Phase_1_total_time = GB_data$GB_3_phase1_total_time_CVC,
                     Phase_1_number_of_times = GB_data$GB_5_phase1_number_of_times_CVC,
```

```
                    Phase_2_total_time = GB_data$GB_7_phase2_total_time_CVC,
                    Phase_2_number_of_times = GB_data$GB_9_phase2_number_of_times_CVC,
                    Phase_2_number_of_times_objects =
                      GB_data$GB_11_phase2_number_of_times_objects_CVC)

print(GB_CVC)
```

```
##    Participant_number Phase_1_total_time Phase_1_number_of_times
## 1                 180                 28                      14
## 2                 122                 13                      13
## 3                 198                 33                      25
## 4                 182                 27                      33
## 5                 132                 40                      34
## 6                 107                 36                      37
## 7                 184                 28                      21
## 8                 137                 38                      32
## 9                 103                 21                      24
## 10                136                 33                      35
## 11                128                 35                      24
## 12                185                 27                      23
## 13                163                 34                      33
## 14                112                 28                      34
## 15                149                 43                      46
## 16                173                 47                      39
## 17                195                 37                      34
## 18                115                 14                      24
## 19                169                 56                      39
## 20                120                 16                      27
## 21                146                 43                      32
## 22                178                 29                      56
## 23                119                 26                      57
## 24                157                 25                      35
##    Phase_2_total_time Phase_2_number_of_times
## 1                  10                      10
## 2                   3                      12
## 3                  22                      10
## 4                  19                      23
## 5                  30                      19
## 6                  32                      31
## 7                  31                      18
## 8                  18                      16
## 9                   1                       6
## 10                 26                      32
## 11                 11                      13
## 12                 33                      17
## 13                 23                      28
## 14                 20                      25
## 15                 26                      22
## 16                 25                      23
## 17                 24                      22
## 18                 18                      16
## 19                 17                      23
## 20                 18                      23
## 21                 38                      22
```

```
## 22                 21                    38
## 23                  3                    14
## 24                 12                    30
##    Phase_2_number_of_times_objects
## 1                                5
## 2                               44
## 3                                3
## 4                                5
## 5                                3
## 6                                6
## 7                                1
## 8                                2
## 9                                4
## 10                               3
## 11                               4
## 12                               4
## 13                               1
## 14                               3
## 15                              22
## 16                               5
## 17                               3
## 18                               2
## 19                               3
## 20                               3
## 21                               0
## 22                               2
## 23                               3
## 24                               2
```

The data table of the GAVC:

```
GB_GAVC <- data.frame(Participant_number = GB_data$Participant_number,
                      Phase_1_total_time = GB_data$GB_4_phase1_total_time_GAVC,
                      Phase_1_number_of_times = GB_data$GB_6_phase1_number_of_times_GAVC,
                      Phase_2_total_time = GB_data$GB_8_phase2_total_time_GAVC,
                      Phase_2_number_of_times = GB_data$GB_10_phase2_number_of_times_GAVC,
                      Phase_2_number_of_times_objects =
                        GB_data$GB_12_phase2_number_of_times_objects_GAVC)

print(GB_GAVC)
```

```
##    Participant_number Phase_1_total_time Phase_1_number_of_times
## 1                 180                 16                      14
## 2                 122                 29                      12
## 3                 198                 38                      25
## 4                 182                 54                      39
## 5                 132                 29                      28
## 6                 107                 36                      34
## 7                 184                 18                      17
## 8                 137                 39                      27
## 9                 103                 29                      20
## 10                136                 42                      32
## 11                128                 28                      29
## 12                185                 27                      27
## 13                163                 33                      34
```

```
## 14                112                 27                    34
## 15                149                 40                    41
## 16                173                 46                    37
## 17                195                 49                    40
## 18                115                 27                    28
## 19                169                 52                    38
## 20                120                 14                    22
## 21                146                 37                    33
## 22                178                 27                    46
## 23                119                  6                    26
## 24                157                 27                    41
##    Phase_2_total_time Phase_2_number_of_times
## 1                  11                      11
## 2                   2                       5
## 3                  21                      15
## 4                  16                      13
## 5                  14                      15
## 6                  27                      22
## 7                  13                      13
## 8                  20                      15
## 9                   1                       3
## 10                 30                      31
## 11                 12                      21
## 12                 15                      19
## 13                 19                      15
## 14                 19                      23
## 15                 18                      21
## 16                 23                      22
## 17                 29                      29
## 18                 16                      15
## 19                 23                      26
## 20                 18                      20
## 21                 28                      30
## 22                 19                      37
## 23                  1                       8
## 24                 10                      16
##    Phase_2_number_of_times_objects
## 1                                4
## 2                                4
## 3                                6
## 4                                2
## 5                                4
## 6                                4
## 7                                2
## 8                                4
## 9                                5
## 10                               3
## 11                               5
## 12                               3
## 13                               4
## 14                               2
## 15                               2
## 16                               6
## 17                               2
```

```
## 18                                5
## 19                                5
## 20                                3
## 21                                4
## 22                                3
## 23                                3
## 24                                1
```

Assumption checking: normal distribution

Phase 1: total gaze duration

The analysis method depends on the normality of the data distribution. This is usually done visually:

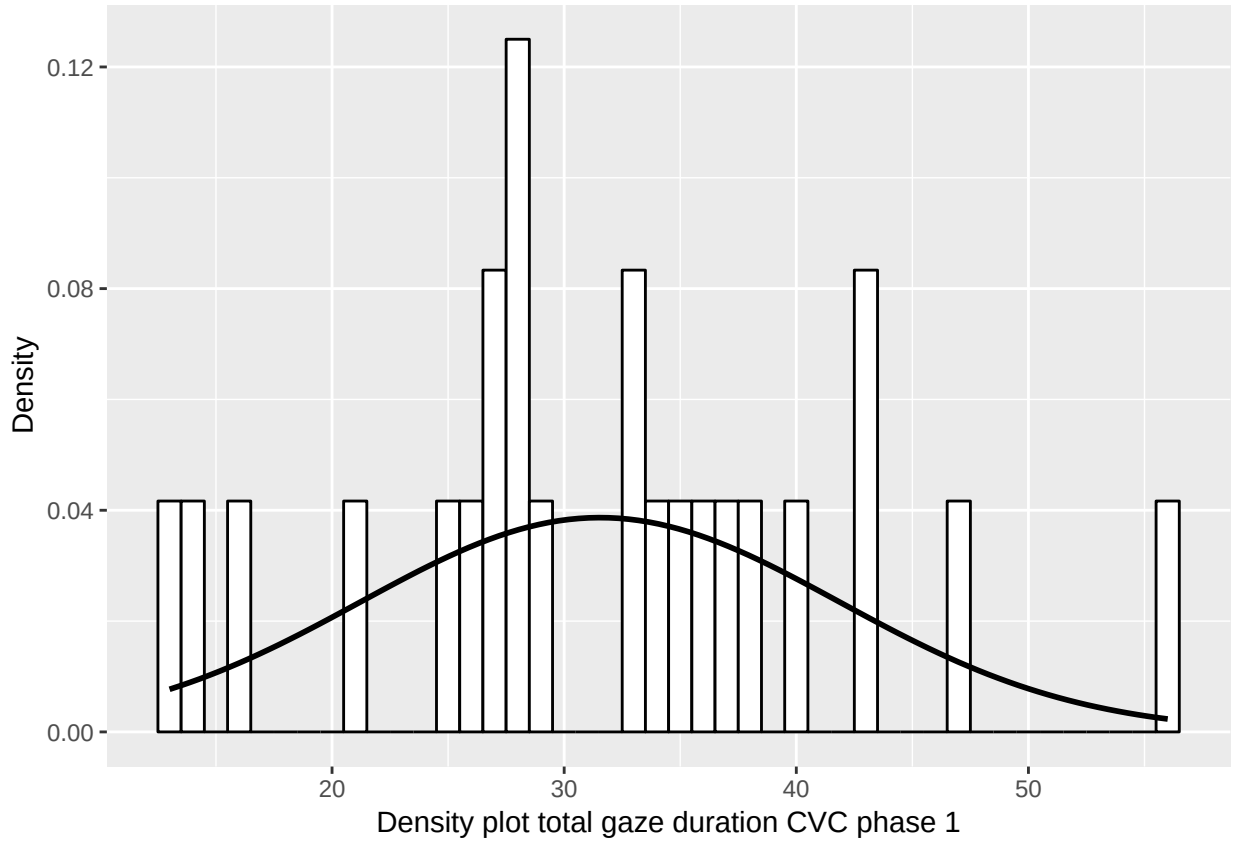
```
stem(GB_CVC$Phase_1_total_time)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   1 | 346
##   2 | 156778889
##   3 | 3345678
##   4 | 0337
##   5 | 6
```

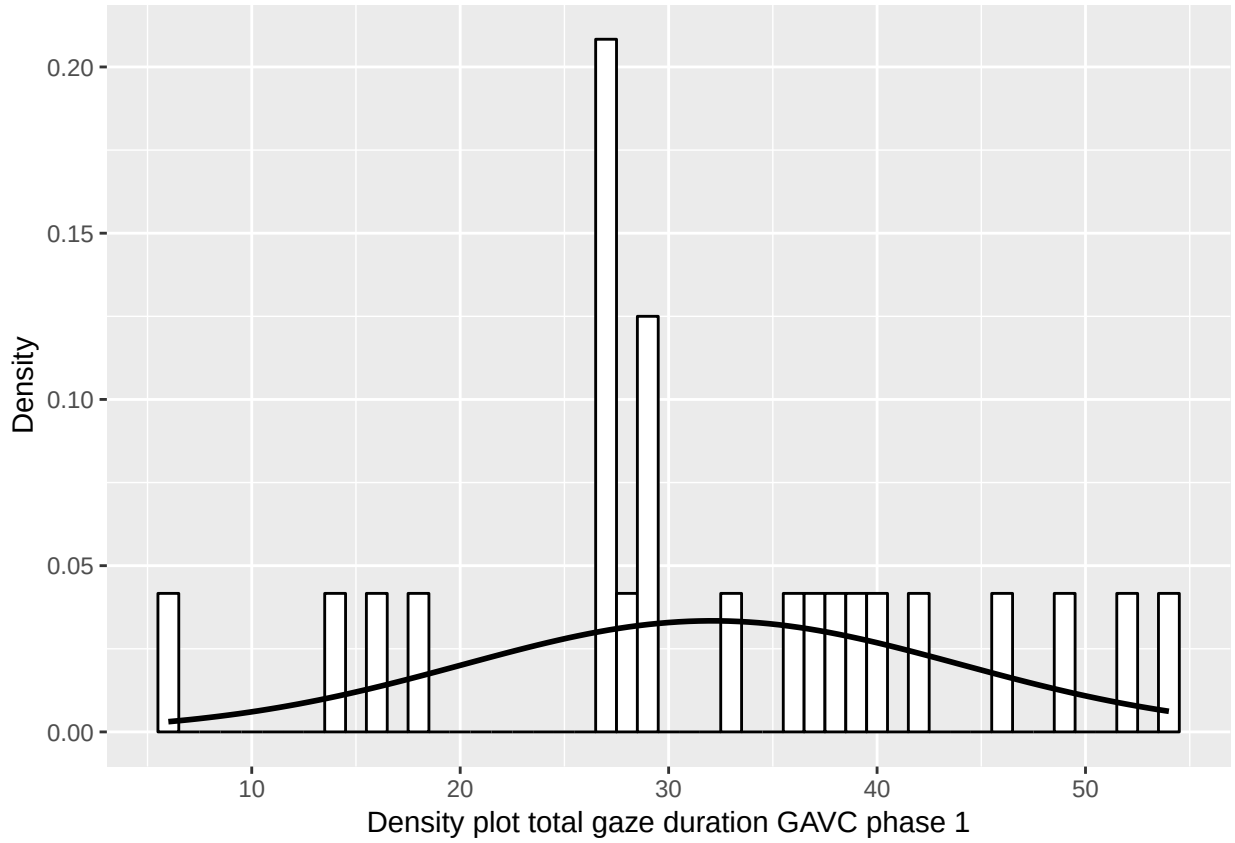
```
stem(GB_GAVC$Phase_1_total_time)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 6
##   1 | 468
##   2 | 777778999
##   3 | 36789
##   4 | 0269
##   5 | 24
```

```
ggplot(GB_CVC,
       aes(Phase_1_total_time)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot total gaze duration CVC phase 1",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_CVC$Phase_1_total_time,
       na.rm=TRUE), sd=sd(GB_CVC$Phase_1_total_time, na.rm=TRUE)), colour="black", size=1)
```

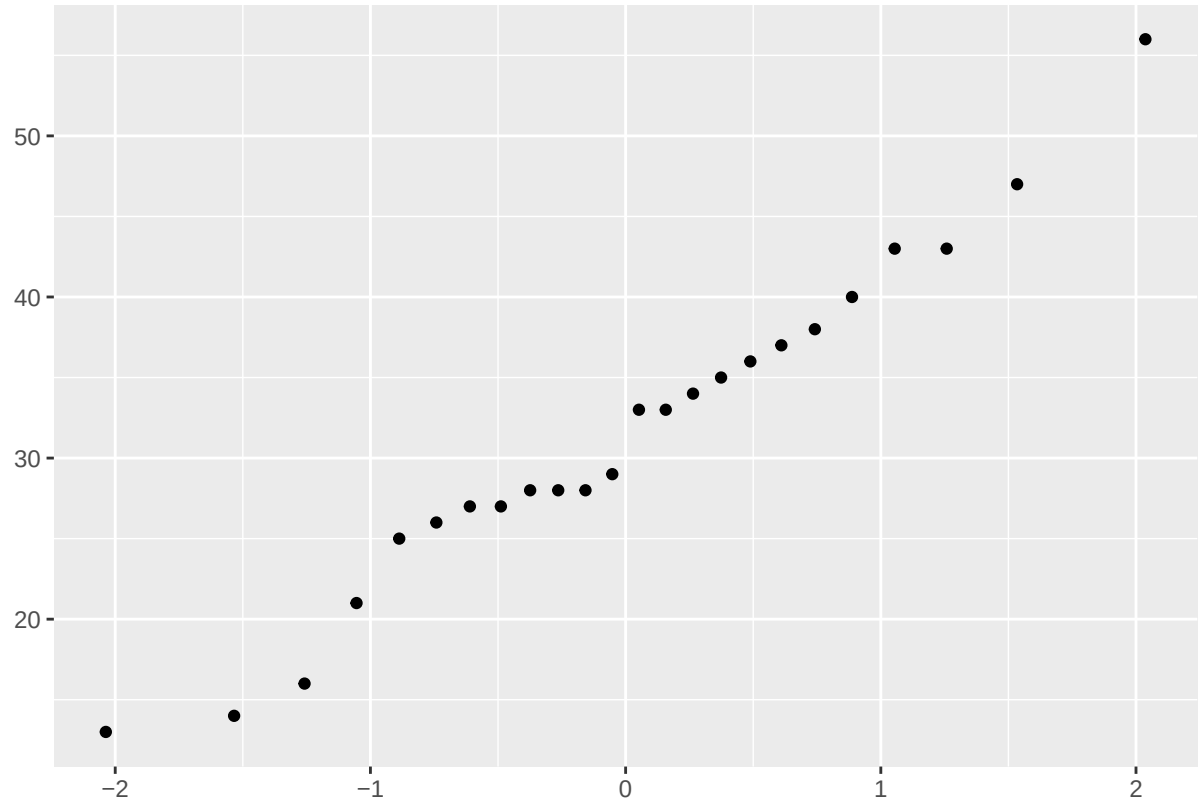



```
ggplot(GB_GAVC,
       aes(Phase_1_total_time)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot total gaze duration GAVC phase 1",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_GAVC$Phase_1_total_time,
       na.rm=TRUE), sd=sd(GB_GAVC$Phase_1_total_time, na.rm=TRUE)), colour="black", size=1)
```




```
qplot(sample=GB_CVC$Phase_1_total_time,
      stat="qq") + labs(x="Q-Q plot total gaze duration CVC phase 1")
```

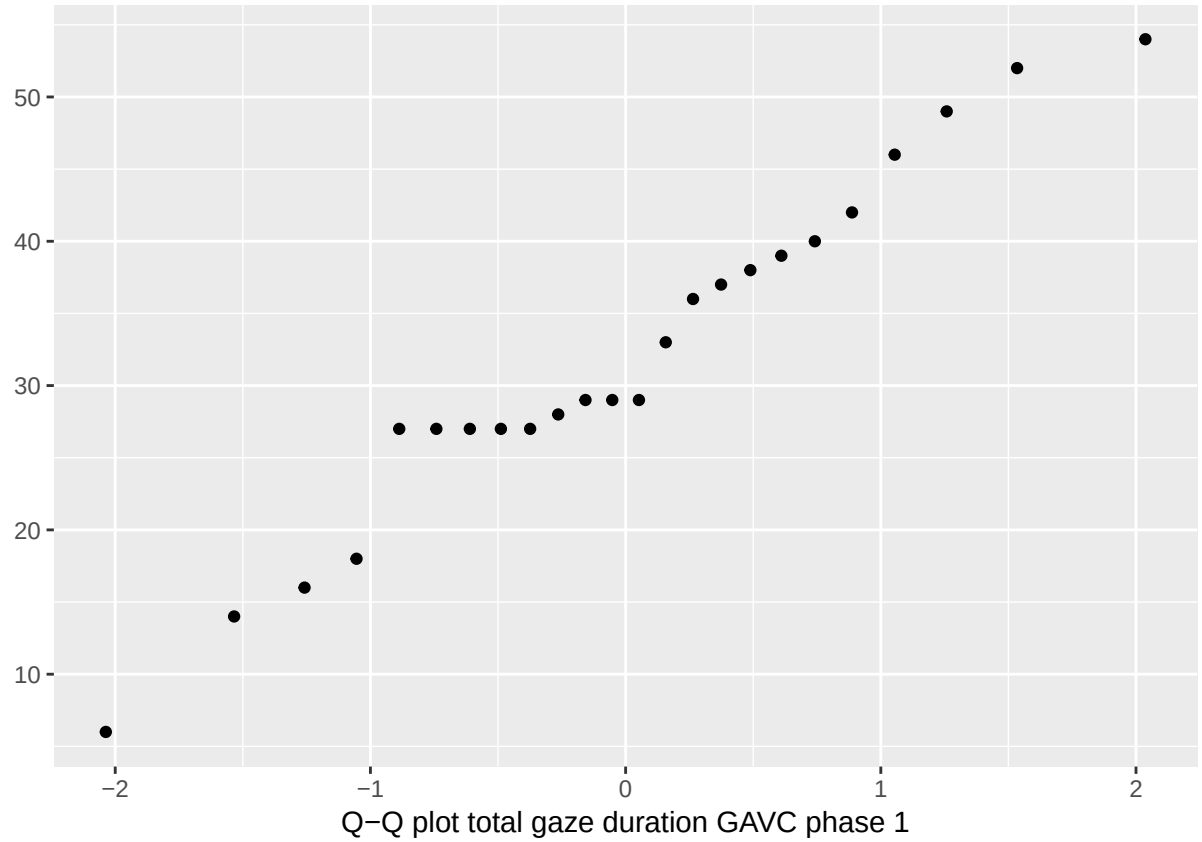
```
## Warning: `stat` is deprecated
```



Q−Q plot total gaze duration CVC phase 1

```
qplot(sample=GB_GAVC$Phase_1_total_time,
      stat="qq") + labs(x="Q-Q plot total gaze duration GAVC phase 1")
```

```
## Warning: `stat` is deprecated
```



By visually inspecting the histogram and qqplot it seems that the data is normally distributed. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(data.frame(GB_CVC$Phase_1_total_time, GB_GAVC$Phase_1_total_time),
                basic = FALSE, norm = TRUE), digits = 3)
```

```
##              GB_CVC.Phase_1_total_time GB_GAVC.Phase_1_total_time
## median                          31.000                     29.000
## mean                            31.542                     32.083
## SE.mean                          2.106                      2.438
## CI.mean.0.95                     4.356                      5.042
## var                            106.433                    142.601
## std.dev                         10.317                     11.942
## coef.var                         0.327                      0.372
## skewness                         0.191                     -0.080
## skew.2SE                         0.202                     -0.085
## kurtosis                        -0.258                     -0.530
## kurt.2SE                        -0.141                     -0.289
## normtest.W                       0.977                      0.969
## normtest.p                       0.830                      0.646
```

The skew.2SE and kurt.2SE are smaller than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis are not significant (at $p < 0.05$). The p-values (indicated by normtest.p) obtained by the Shapiro-Wilk test are >0.05 and thus both the CVC and GAVC are normally distributed.

Phase 1: number of gazes

The analysis method depends on the normality of the data distribution. This is usually done visually:

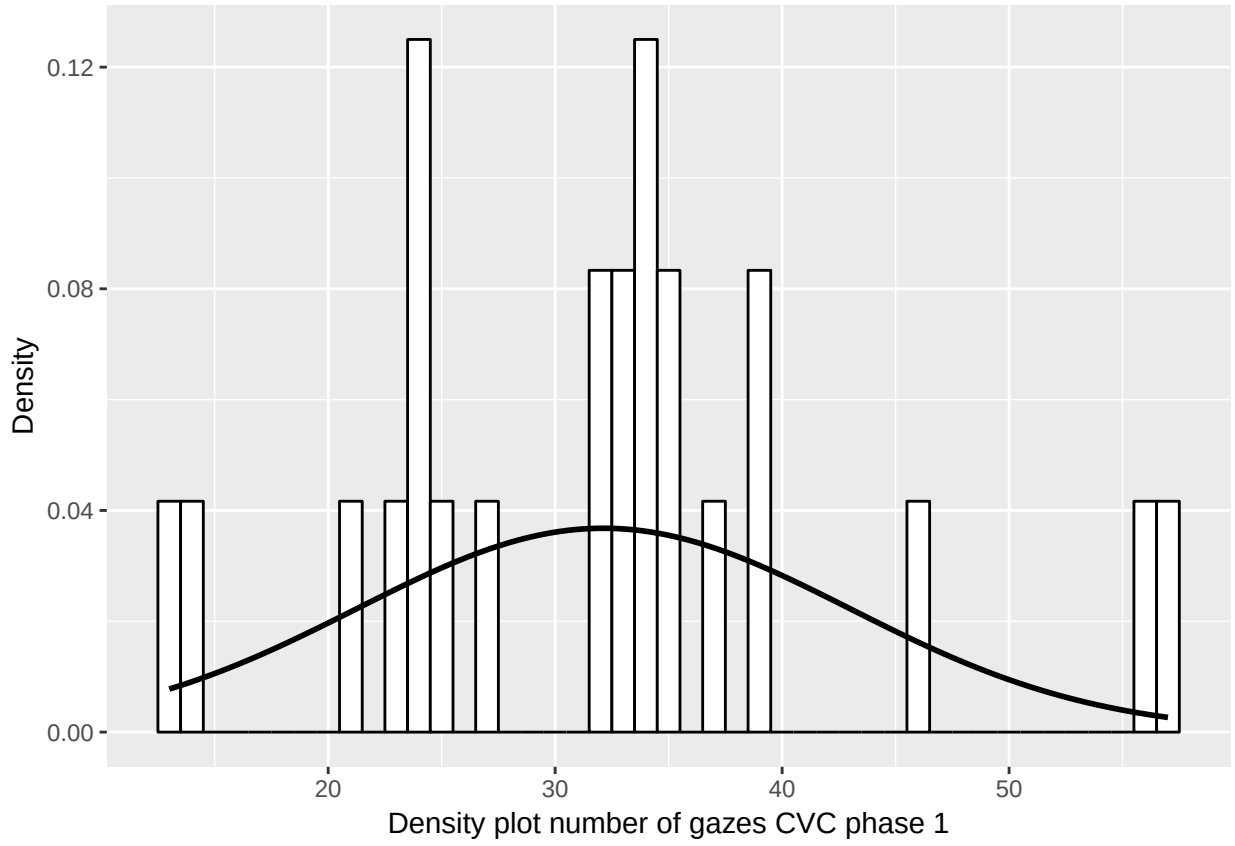
```
stem(GB_CVC$Phase_1_number_of_times)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   1 | 34
##   2 | 1344457
##   3 | 223344455799
##   4 | 6
##   5 | 67
```

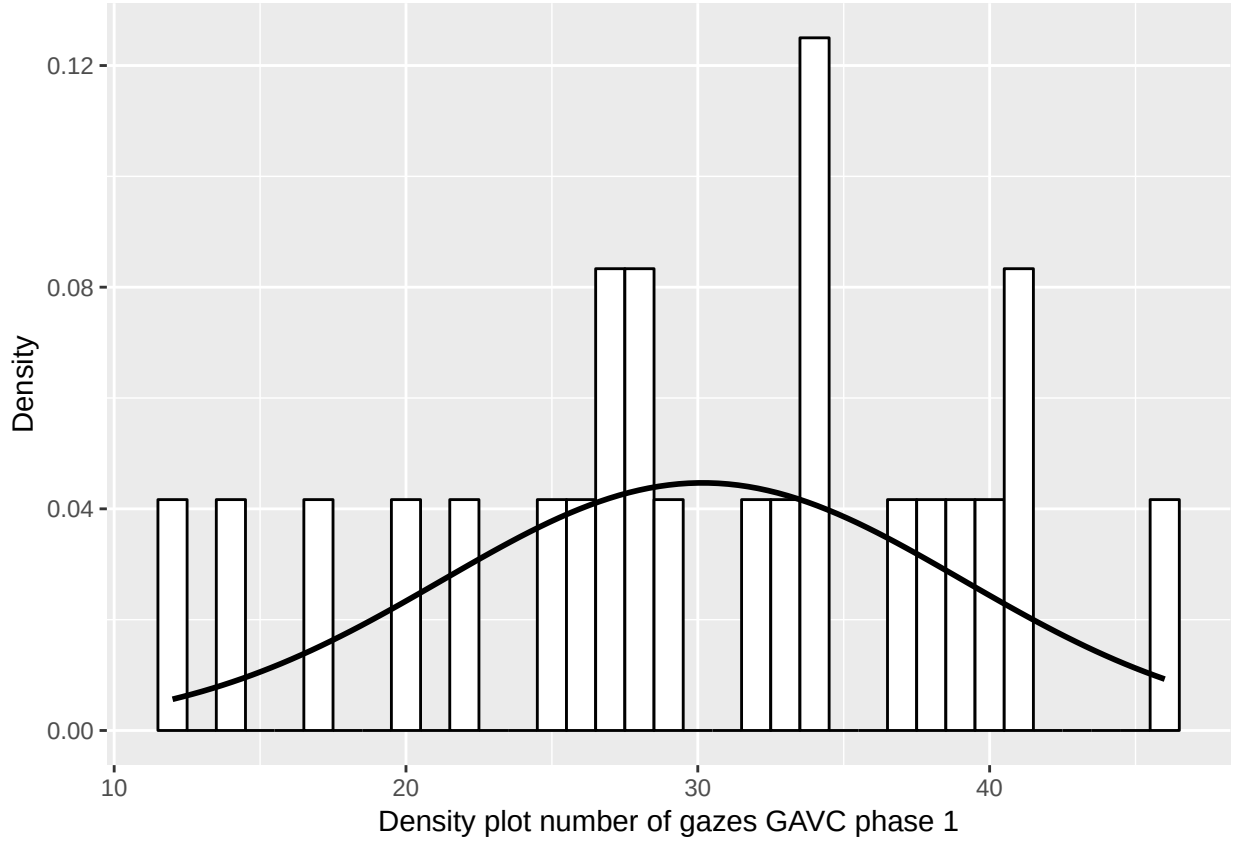
```
stem(GB_GAVC$Phase_1_number_of_times)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   1 | 247
##   2 | 025677889
##   3 | 23444789
##   4 | 0116
```

```
ggplot(GB_CVC,
       aes(Phase_1_number_of_times)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot number of gazes CVC phase 1",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_CVC$Phase_1_number_of_times,
       na.rm=TRUE), sd=sd(GB_CVC$Phase_1_number_of_times, na.rm=TRUE)), colour="black", size=1)
```

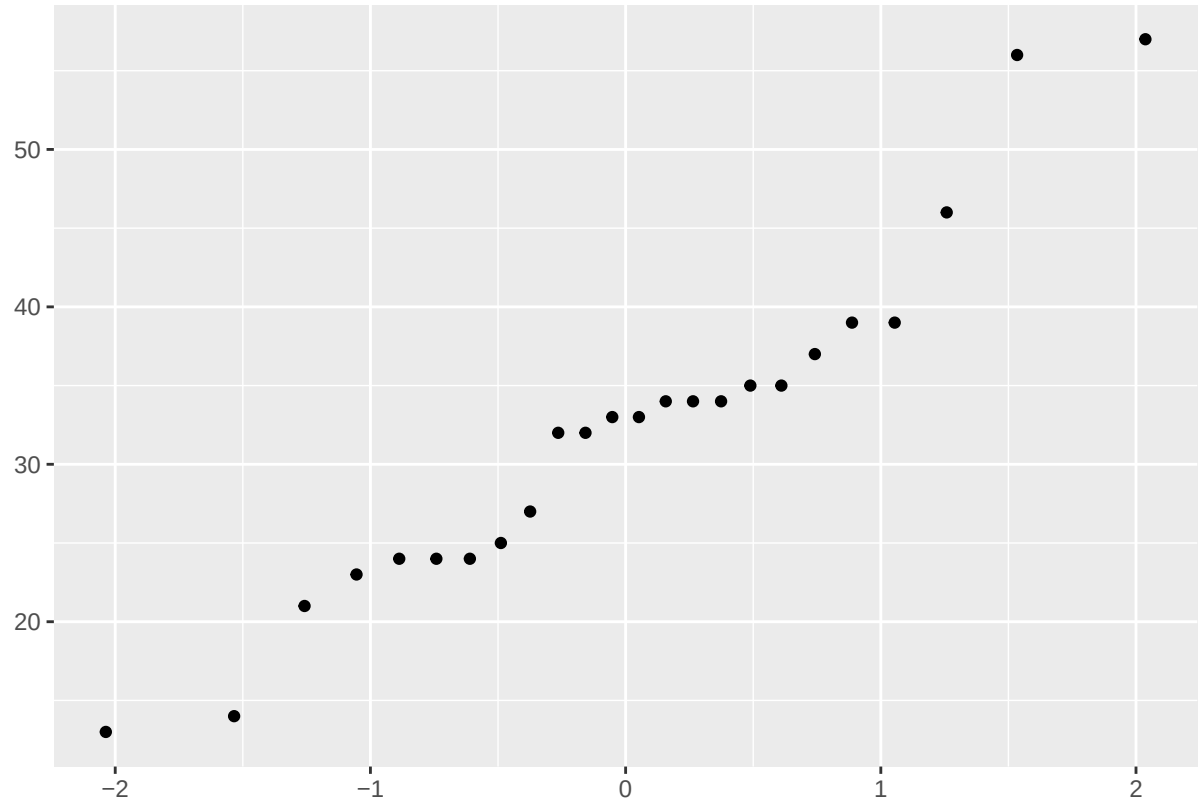


```
ggplot(GB_GAVC,
       aes(Phase_1_number_of_times)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot number of gazes GAVC phase 1",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_GAVC$Phase_1_number_of_times,
       na.rm=TRUE), sd=sd(GB_GAVC$Phase_1_number_of_times, na.rm=TRUE)), colour="black", size=1)
```



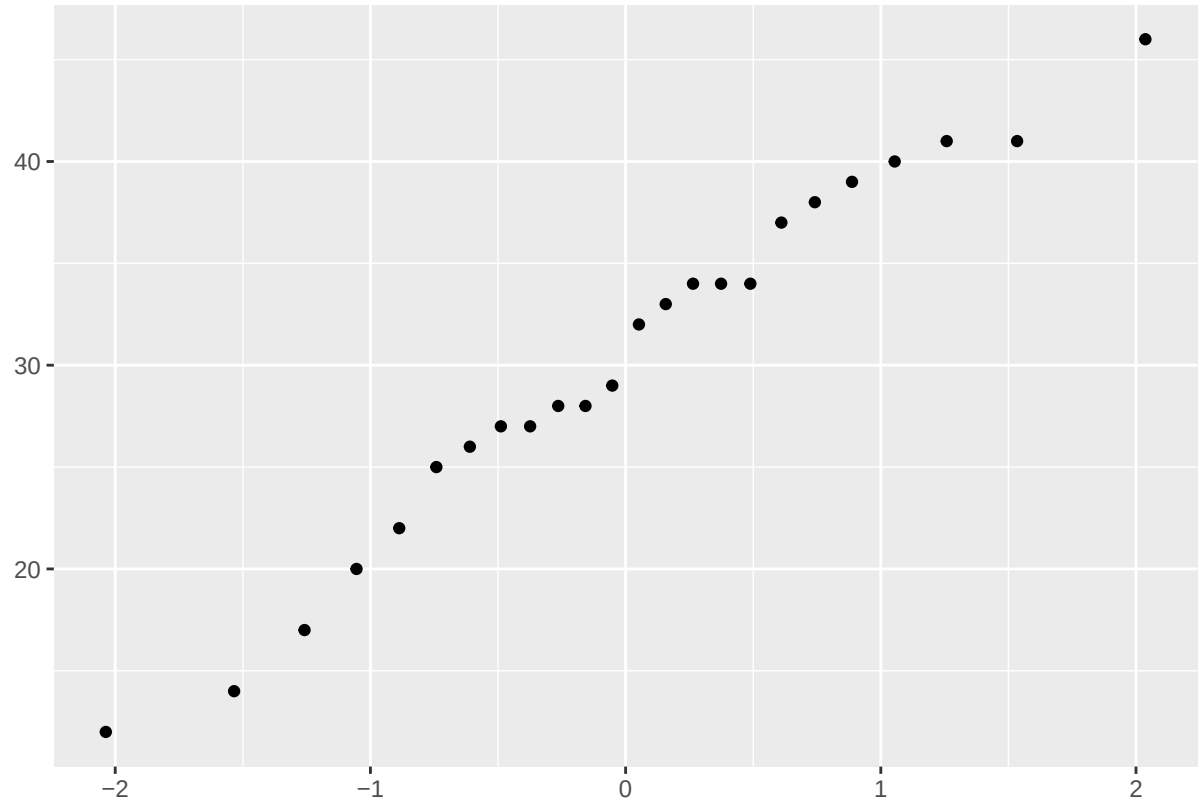

```
qplot(sample=GB_CVC$Phase_1_number_of_times,
      stat="qq") + labs(x="Q-Q plot number of gazes CVC phase 1")
```

```
## Warning: `stat` is deprecated
```




```
qplot(sample=GB_GAVC$Phase_1_number_of_times,
      stat="qq") + labs(x="Q-Q plot number of gazes GAVC phase 1")
```

```
## Warning: `stat` is deprecated
```



Q−Q plot number of gazes GAVC phase 1

By visually inspecting the histogram and qqplot it seems that the data is normally distributed. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(data.frame(GB_CVC$Phase_1_number_of_times, GB_GAVC$Phase_1_number_of_times),
                basic = FALSE, norm = TRUE), digits = 3)
```

```
##              GB_CVC.Phase_1_number_of_times
## median                               33.000
## mean                                 32.125
## SE.mean                               2.214
## CI.mean.0.95                          4.579
## var                                 117.592
## std.dev                              10.844
## coef.var                              0.338
## skewness                              0.500
## skew.2SE                              0.530
## kurtosis                              0.105
## kurt.2SE                              0.057
## normtest.W                            0.939
## normtest.p                            0.153
##              GB_GAVC.Phase_1_number_of_times
## median                               30.500
## mean                                 30.167
## SE.mean                               1.822
## CI.mean.0.95                          3.770
## var                                  79.710
## std.dev                               8.928
```

```
## coef.var                              0.296
## skewness                             -0.307
## skew.2SE                             -0.325
## kurtosis                             -0.795
## kurt.2SE                             -0.433
## normtest.W                            0.973
## normtest.p                            0.731
```

The skew.2SE and kurt.2SE are smaller than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis are not significant (at $p < 0.05$). The p-values (indicated by normtest.p) obtained by the Shapiro-Wilk test are >0.05 and thus both the CVC and GAVC are normally distributed.

Phase 2: total gaze duration

The analysis method depends on the normality of the data distribution. This is usually done visually:

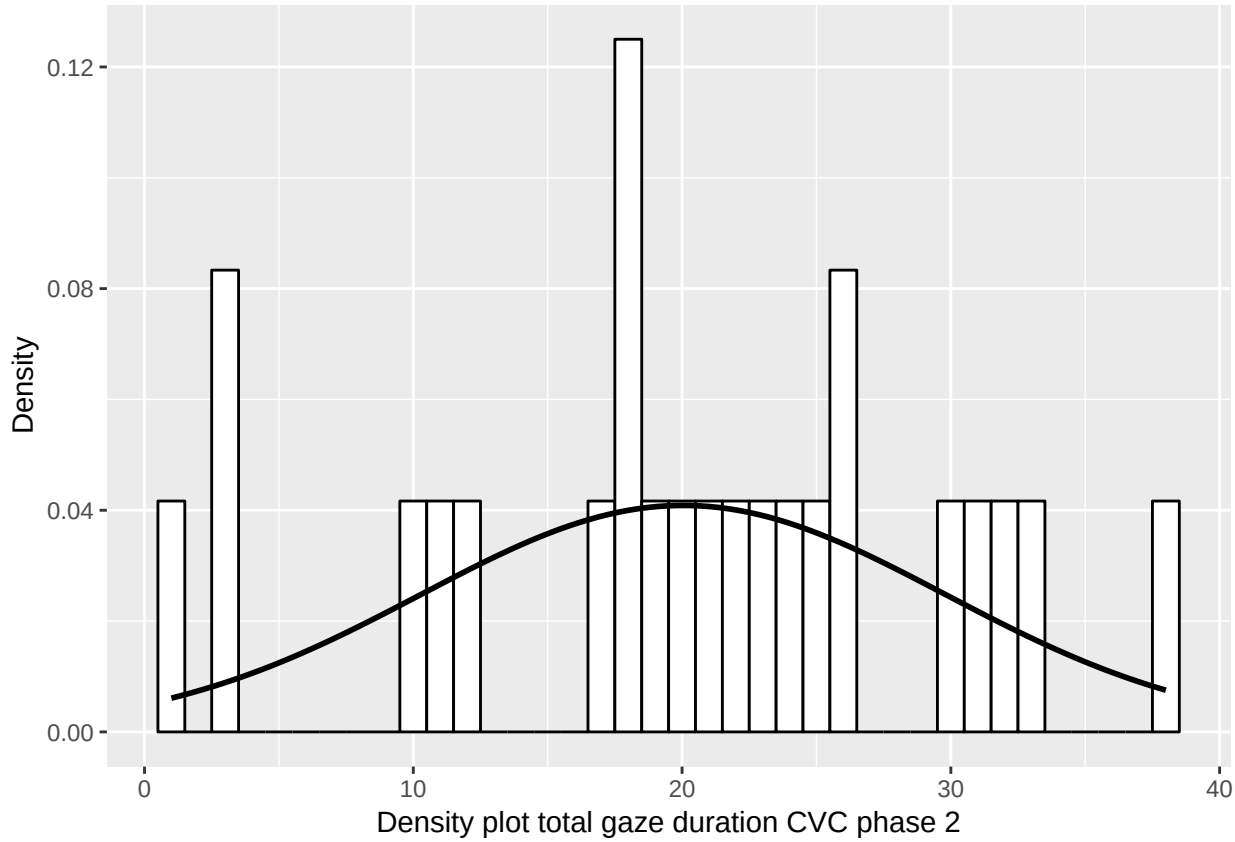
```
stem(GB_CVC$Phase_2_total_time)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 133
##   1 | 01278889
##   2 | 01234566
##   3 | 01238
```

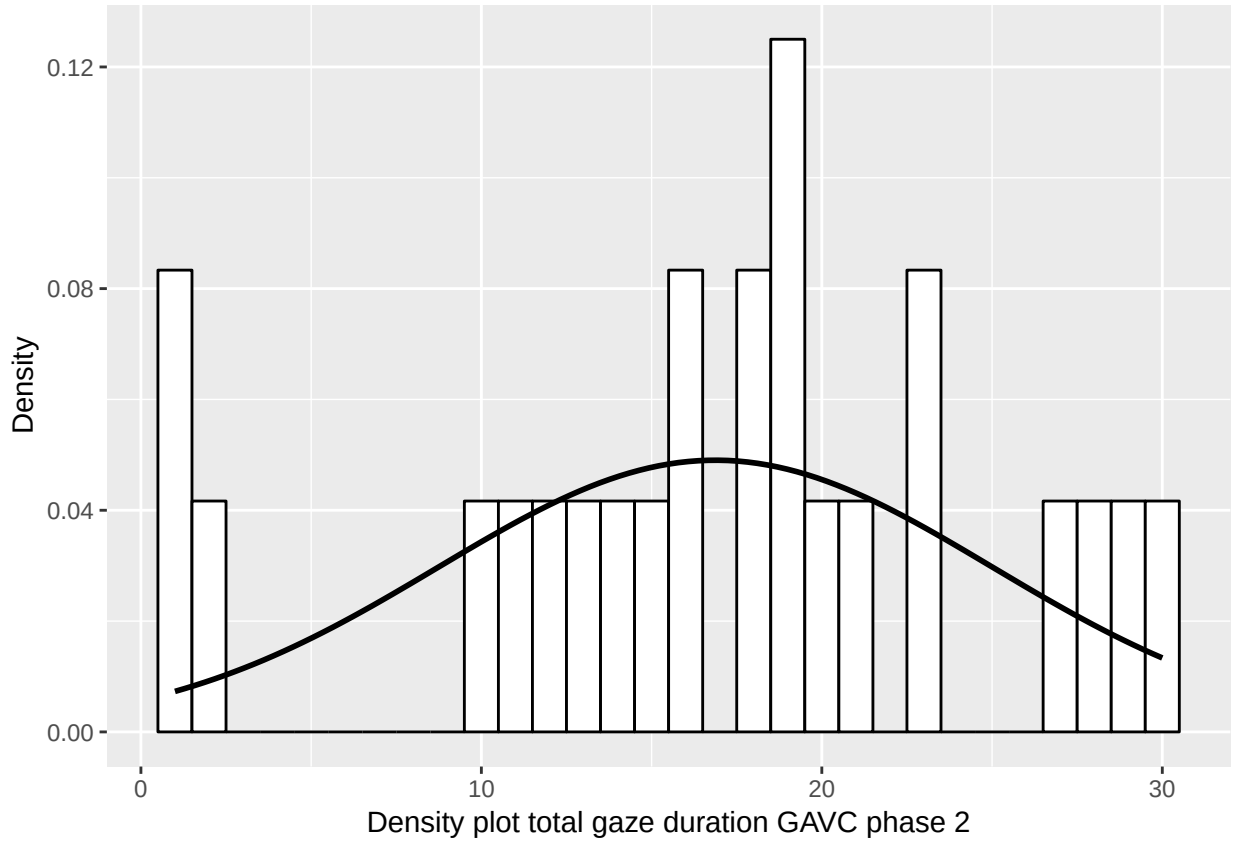
```
stem(GB_GAVC$Phase_2_total_time)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 112
##   1 | 0123456688999
##   2 | 0133789
##   3 | 0
```

```
ggplot(GB_CVC,
       aes(Phase_2_total_time)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot total gaze duration CVC phase 2",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_CVC$Phase_2_total_time,
       na.rm=TRUE), sd=sd(GB_CVC$Phase_2_total_time, na.rm=TRUE)), colour="black", size=1)
```

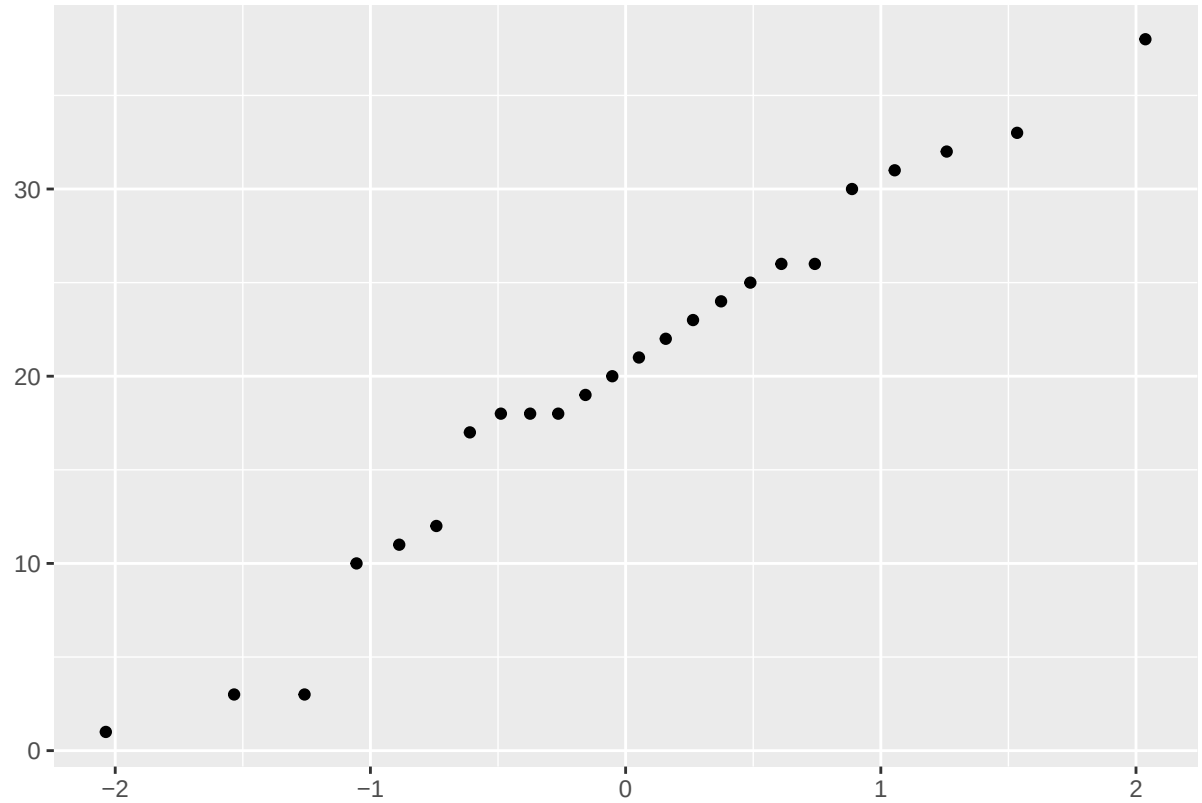



```
ggplot(GB_GAVC,
       aes(Phase_2_total_time)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot total gaze duration GAVC phase 2",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_GAVC$Phase_2_total_time,
       na.rm=TRUE), sd=sd(GB_GAVC$Phase_2_total_time, na.rm=TRUE)), colour="black", size=1)
```



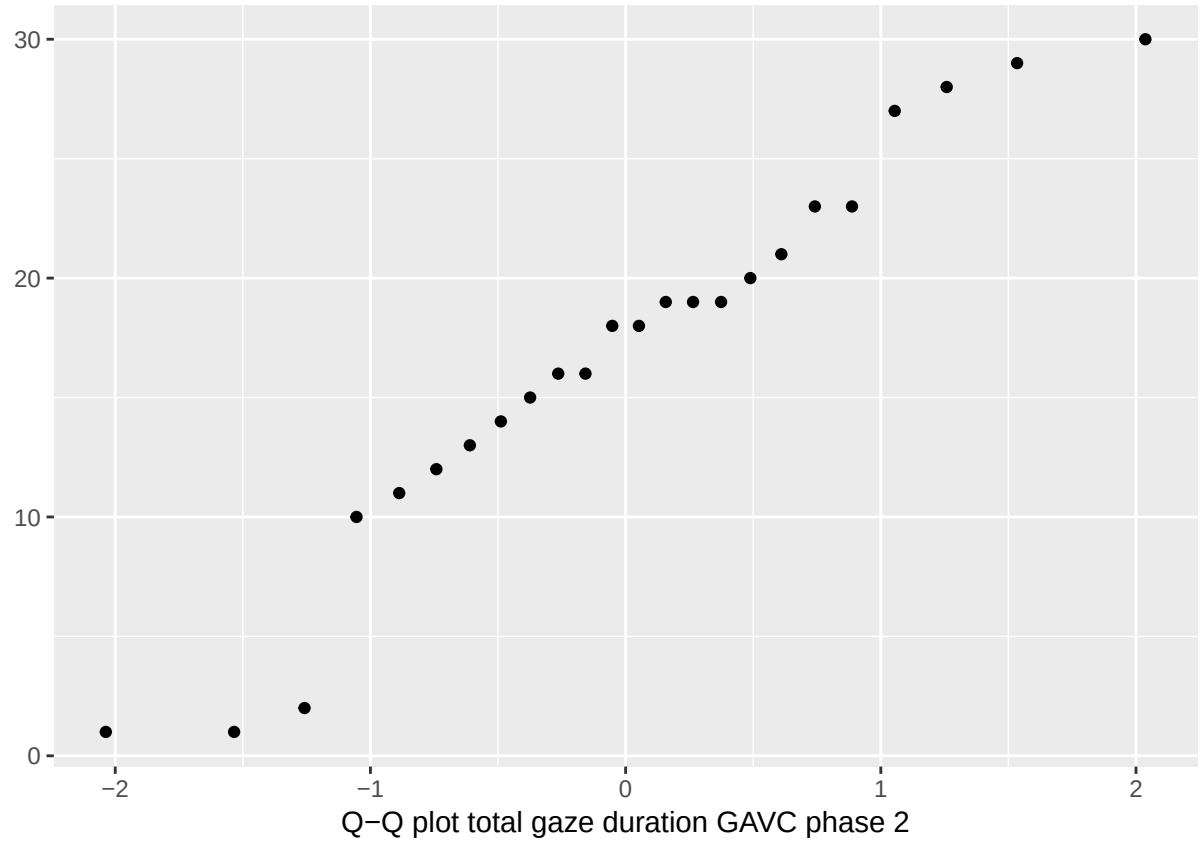

```
qplot(sample=GB_CVC$Phase_2_total_time,
      stat="qq") + labs(x="Q-Q plot total gaze duration CVC phase 2")
```

```
## Warning: `stat` is deprecated
```




```
qplot(sample=GB_GAVC$Phase_2_total_time,
      stat="qq") + labs(x="Q-Q plot total gaze duration GAVC phase 2")
```

```
## Warning: `stat` is deprecated
```



By visually inspecting the histogram and qqplot it seems that the data is normally distributed. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(data.frame(GB_CVC$Phase_2_total_time, GB_GAVC$Phase_2_total_time),
                basic = FALSE, norm = TRUE), digits = 3)
```

```
##              GB_CVC.Phase_2_total_time GB_GAVC.Phase_2_total_time
## median                          20.500                     18.000
## mean                            20.042                     16.875
## SE.mean                          1.992                      1.661
## CI.mean.0.95                     4.121                      3.436
## var                             95.259                     66.201
## std.dev                          9.760                      8.136
## coef.var                         0.487                      0.482
## skewness                        -0.292                     -0.385
## skew.2SE                        -0.309                     -0.408
## kurtosis                        -0.696                     -0.522
## kurt.2SE                        -0.379                     -0.284
## normtest.W                       0.968                      0.949
## normtest.p                       0.611                      0.252
```

The skew.2SE and kurt.2SE are smaller than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis are not significant (at $p < 0.05$). The p-values (indicated by normtest.p) obtained by the Shapiro-Wilk test are >0.05 and thus both the CVC and GAVC are normally distributed.

Phase 2: number of gazes

The analysis method depends on the normality of the data distribution. This is usually done visually:

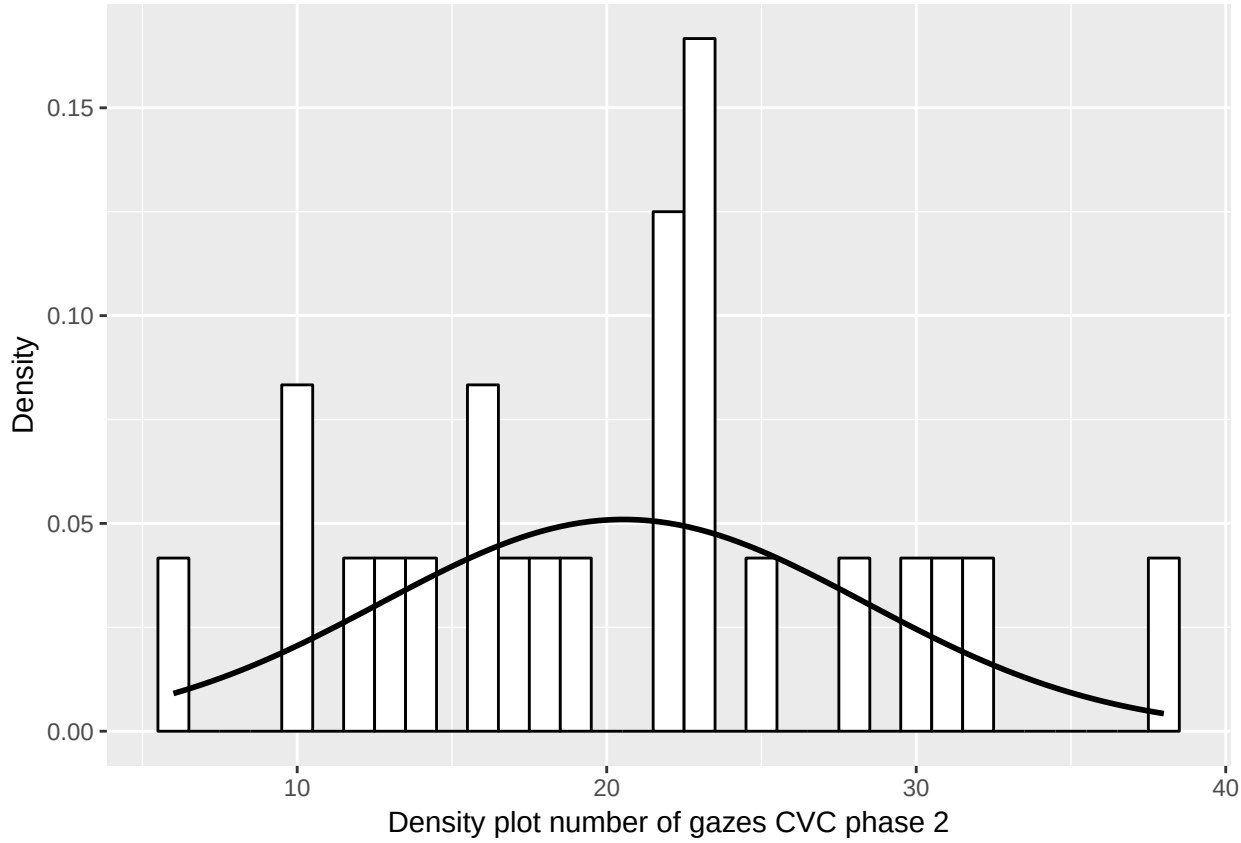
```
stem(GB_CVC$Phase_2_number_of_times)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 6
##   1 | 0023466789
##   2 | 222333358
##   3 | 0128
```

```
stem(GB_GAVC$Phase_2_number_of_times)
```

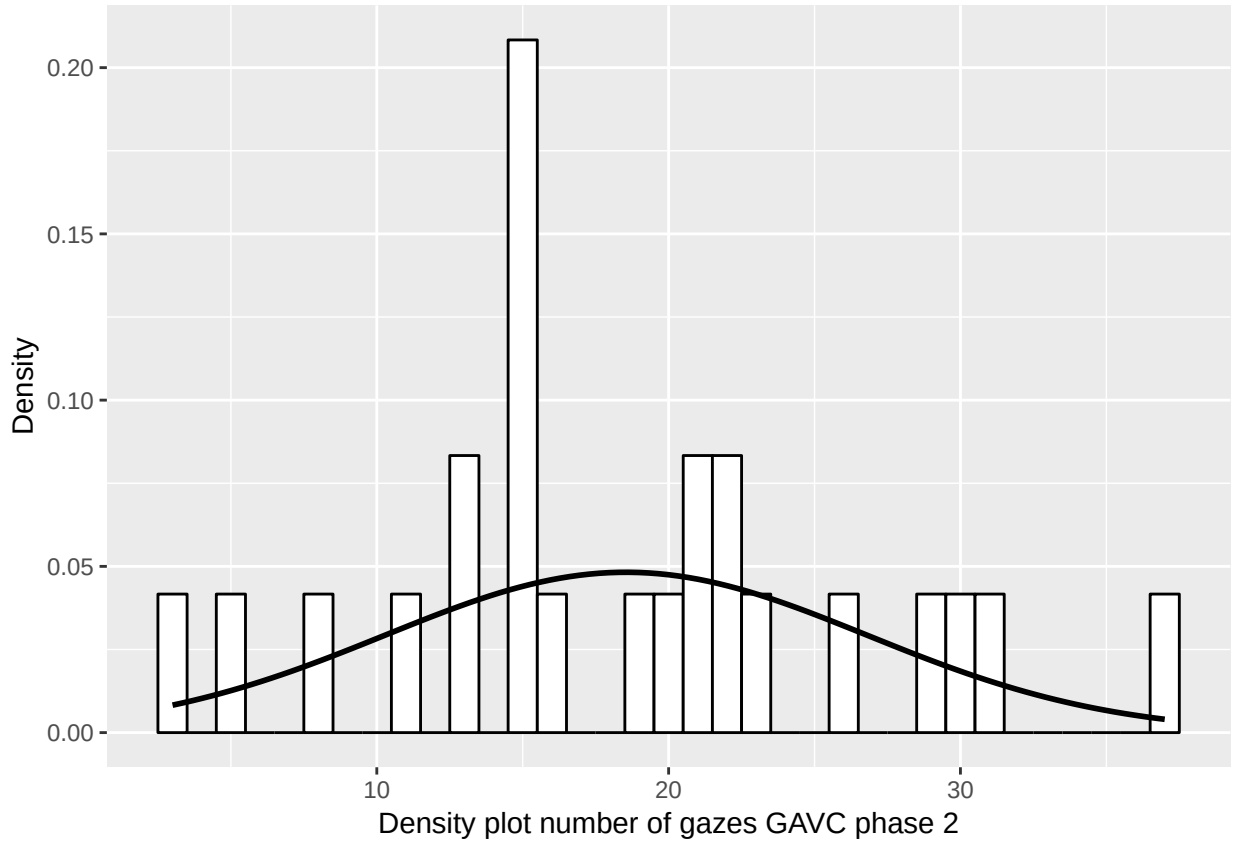
```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 358
##   1 | 1335555569
##   2 | 01122369
##   3 | 017
```

```
ggplot(GB_CVC,
       aes(Phase_2_number_of_times)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot number of gazes CVC phase 2",
       y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_CVC$Phase_2_number_of_times,
       na.rm=TRUE), sd=sd(GB_CVC$Phase_2_number_of_times, na.rm=TRUE)), colour="black", size=1)
```



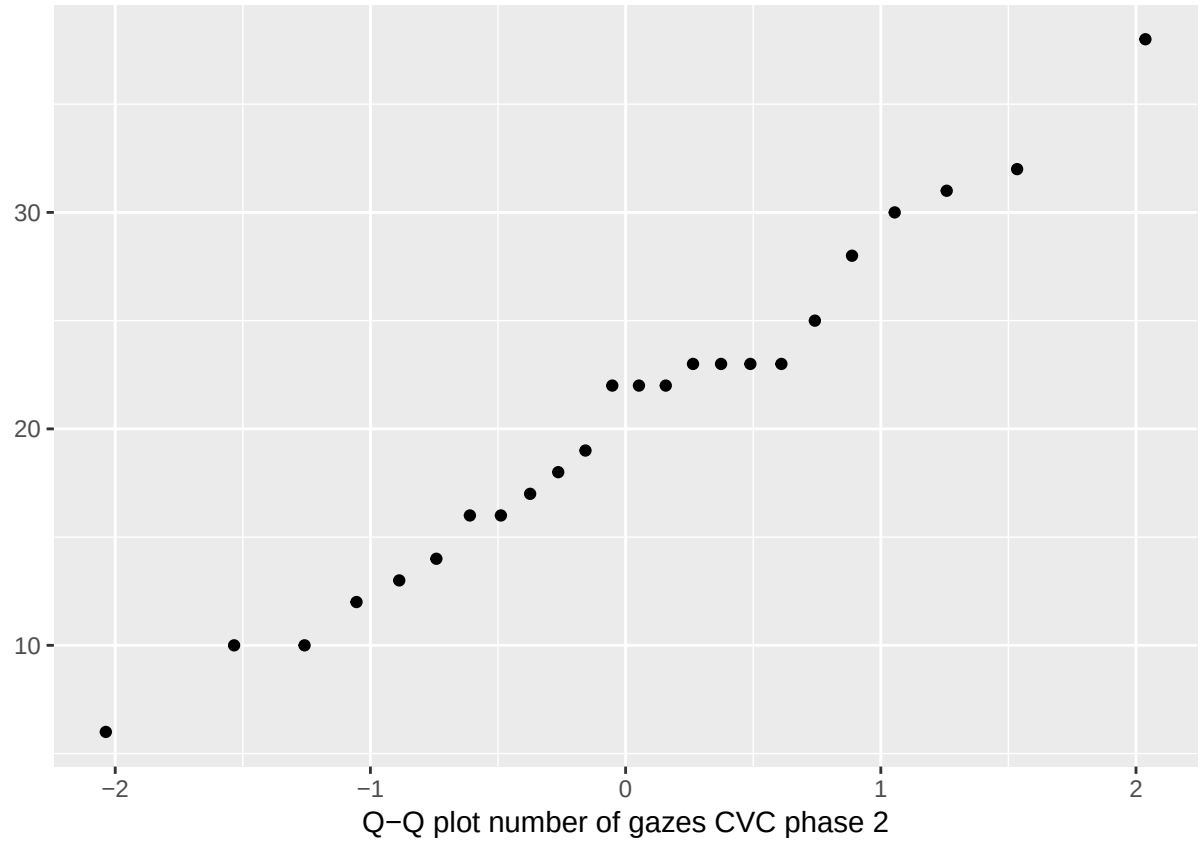

```
ggplot(GB_GAVC,
```

```
      aes(Phase_2_number_of_times)) + geom_histogram(aes(y=..density..), binwidth = 1,
      colour="black", fill="white") + labs(x="Density plot number of gazes GAVC phase 2",
      y="Density") + stat_function(fun=dnorm, args=list(mean=mean(GB_GAVC$Phase_2_number_of_times,
      na.rm=TRUE), sd=sd(GB_GAVC$Phase_2_number_of_times, na.rm=TRUE)), colour="black", size=1)
```



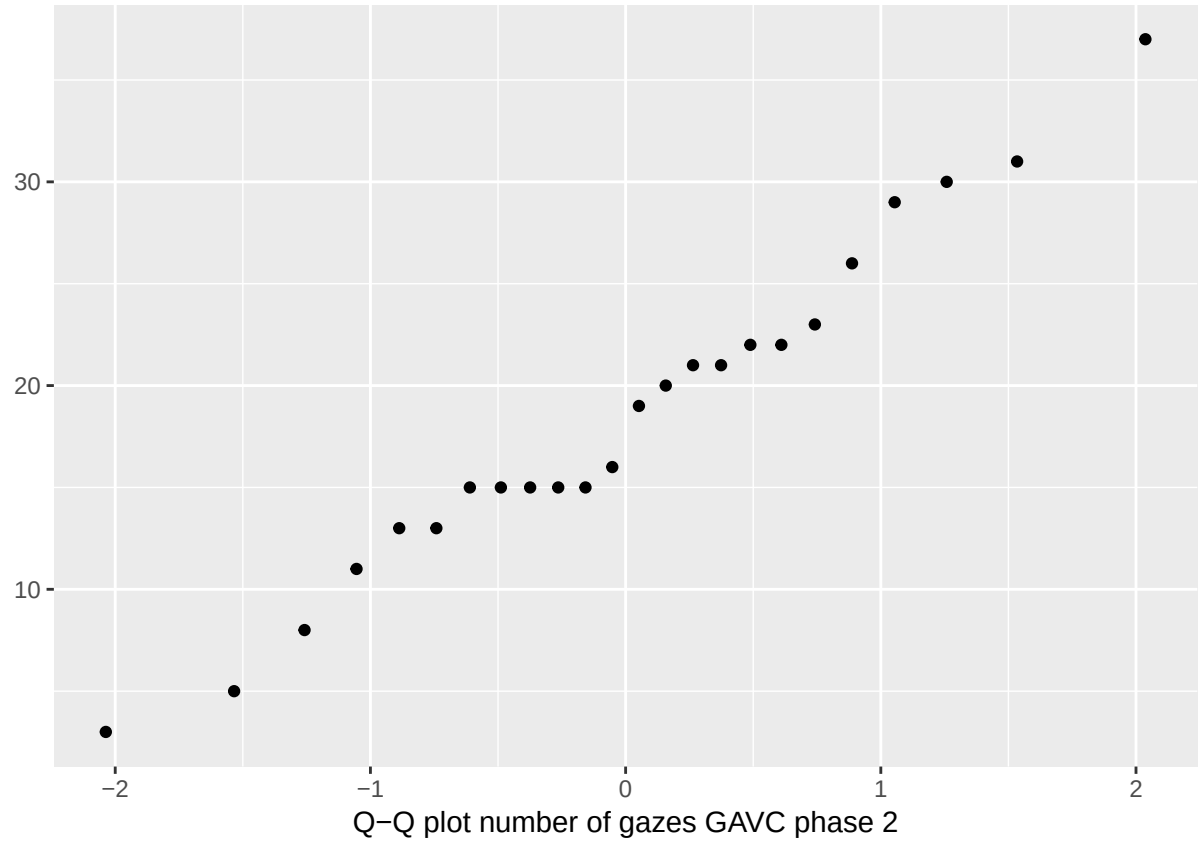

```
qplot(sample=GB_CVC$Phase_2_number_of_times,
      stat="qq") + labs(x="Q-Q plot number of gazes CVC phase 2")
```

```
## Warning: `stat` is deprecated
```




```
qplot(sample=GB_GAVC$Phase_2_number_of_times,
      stat="qq") + labs(x="Q-Q plot number of gazes GAVC phase 2")
```

```
## Warning: `stat` is deprecated
```



By visually inspecting the histogram and qqplot it seems that the data is normally distributed. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(data.frame(GB_CVC$Phase_2_number_of_times, GB_GAVC$Phase_2_number_of_times),
                basic = FALSE, norm = TRUE), digits = 3)
```

```
##              GB_CVC.Phase_2_number_of_times
## median                               22.000
## mean                                 20.542
## SE.mean                               1.598
## CI.mean.0.95                          3.306
## var                                  61.303
## std.dev                               7.830
## coef.var                              0.381
## skewness                              0.207
## skew.2SE                              0.219
## kurtosis                             -0.621
## kurt.2SE                             -0.338
## normtest.W                            0.981
## normtest.p                            0.908
##              GB_GAVC.Phase_2_number_of_times
## median                               17.500
## mean                                 18.542
## SE.mean                               1.689
## CI.mean.0.95                          3.493
## var                                  68.433
## std.dev                               8.272
```

```
## coef.var                              0.446
## skewness                              0.226
## skew.2SE                              0.240
## kurtosis                             -0.478
## kurt.2SE                             -0.260
## normtest.W                            0.978
## normtest.p                            0.847
```

The skew.2SE and kurt.2SE are smaller than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis are not significant (at $p < 0.05$). The p-values (indicated by normtest.p) obtained by the Shapiro-Wilk test are >0.05 and thus both the CVC and GAVC are normally distributed.

Phase 2: number of gazes followed to objects

The analysis method depends on the normality of the data distribution. This is usually done visually:

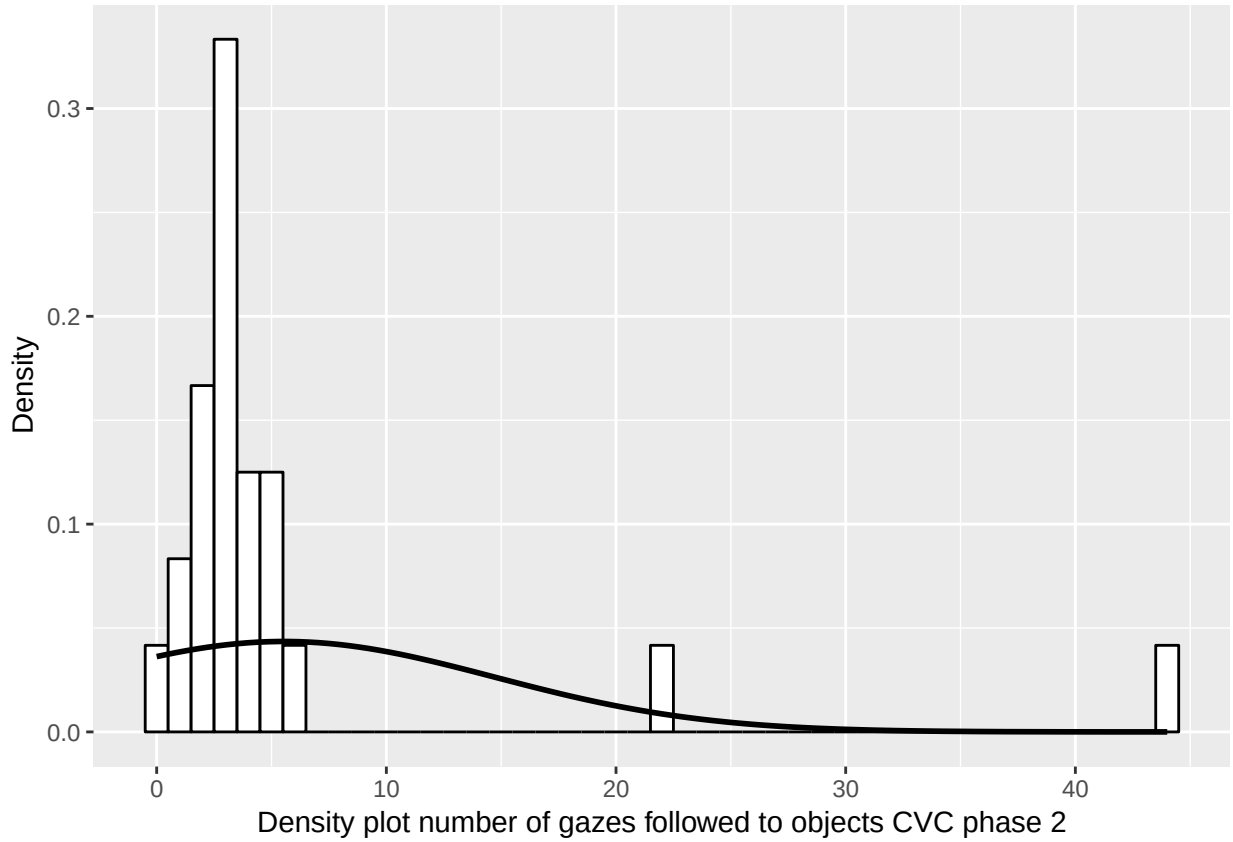
```
stem(GB_CVC$Phase_2_number_of_times_objects)
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   0 | 0112222333333334445556
##   1 |
##   2 | 2
##   3 |
##   4 | 4
```

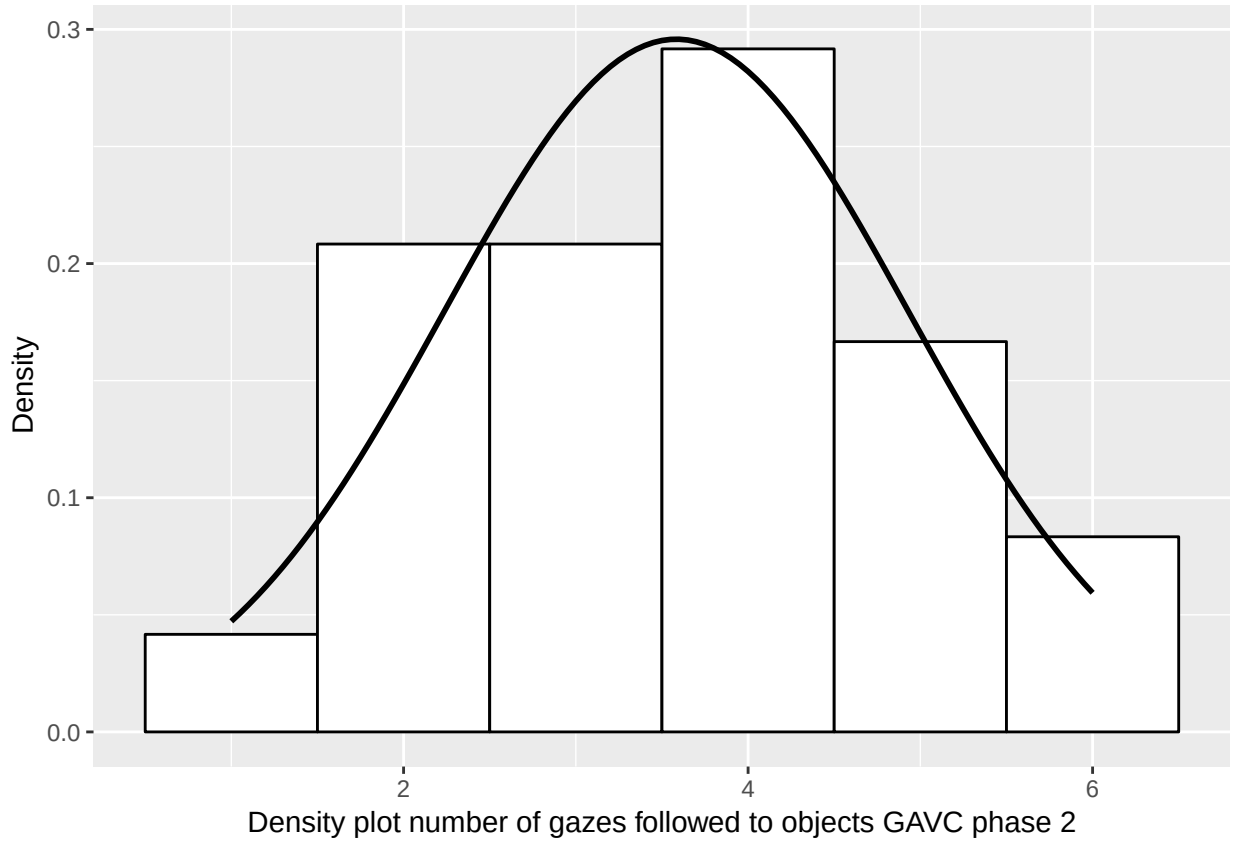
```
stem(GB_GAVC$Phase_2_number_of_times_objects)
```

```
##
##   The decimal point is at the |
##
##   0 | 0
##   2 | 0000000000
##   4 | 00000000000
##   6 | 00
```

```
ggplot(GB_CVC,
       aes(Phase_2_number_of_times_objects)) + geom_histogram(aes(y=..density..),
       binwidth = 1, colour="black",
       fill="white") + labs(x="Density plot number of gazes followed to objects CVC phase 2",
       y="Density") + stat_function(fun=dnorm,
       args=list(mean=mean(GB_CVC$Phase_2_number_of_times_objects, na.rm=TRUE),
       sd=sd(GB_CVC$Phase_2_number_of_times_objects, na.rm=TRUE)), colour="black", size=1)
```

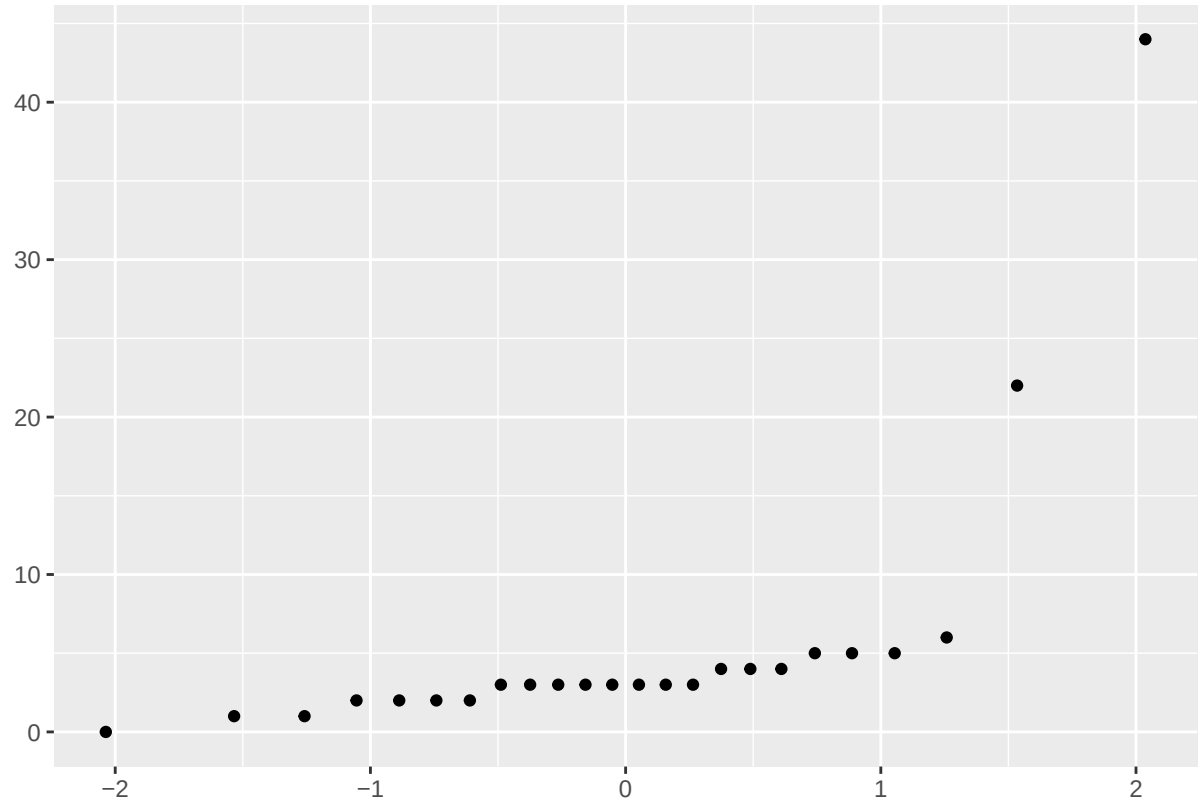



```
ggplot(GB_GAVC,
       aes(Phase_2_number_of_times_objects)) + geom_histogram(aes(y=..density..),
       binwidth = 1, colour="black",
       fill="white") + labs(x="Density plot number of gazes followed to objects GAVC phase 2",
       y="Density") + stat_function(fun=dnorm,
       args=list(mean=mean(GB_GAVC$Phase_2_number_of_times_objects, na.rm=TRUE),
       sd=sd(GB_GAVC$Phase_2_number_of_times_objects, na.rm=TRUE)), colour="black", size=1)
```



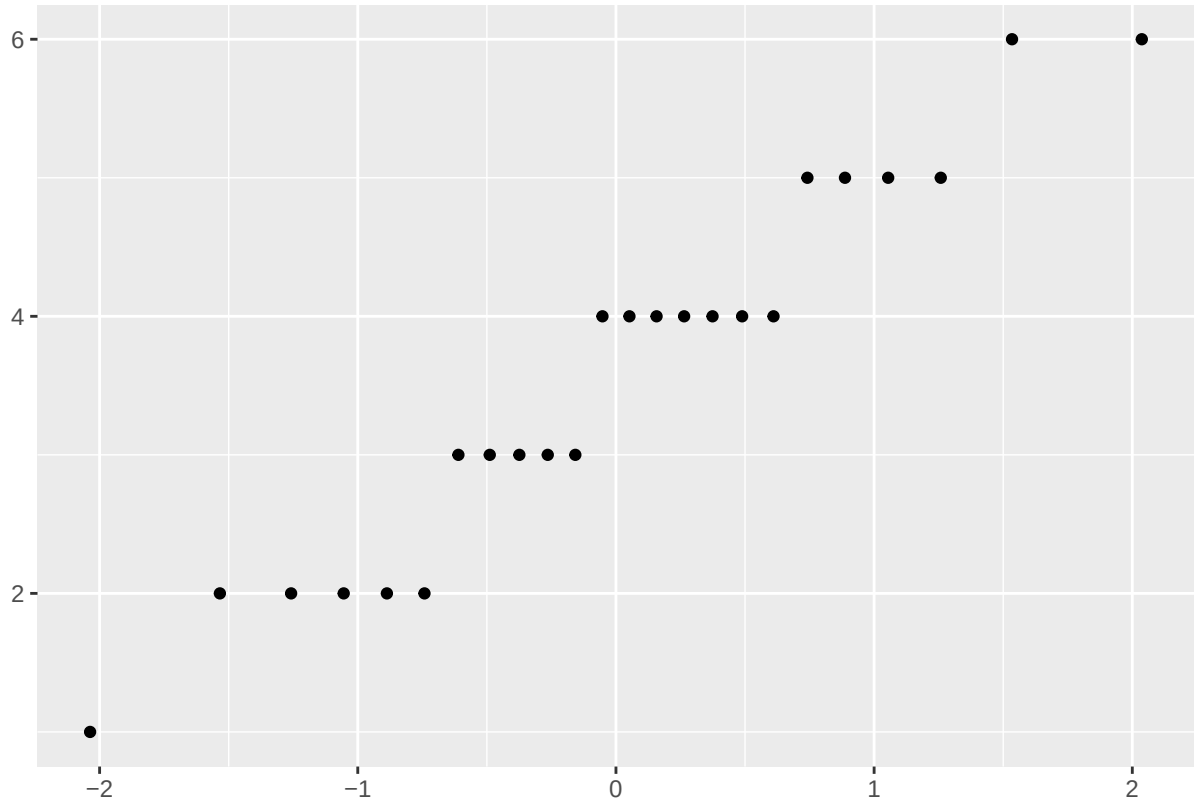

```
qplot(sample=GB_CVC$Phase_2_number_of_times_objects,
      stat="qq") + labs(x="Q-Q plot number of gazes followed to objects CVC phase 2")
```

```
## Warning: `stat` is deprecated
```




```
qplot(sample=GB_GAVC$Phase_2_number_of_times_objects,
      stat="qq") + labs(x="Q-Q plot number of gazes followed to objects GAVC phase 2")
```

```
## Warning: `stat` is deprecated
```



Q−Q plot number of gazes followed to objects GAVC phase 2

By visually inspecting the histogram and qqplot it seems that the data is not normally distributed at all. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(data.frame(GB_CVC$Phase_2_number_of_times_objects,
                           GB_GAVC$Phase_2_number_of_times_objects),
                           basic = FALSE, norm = TRUE), digits = 3)
```

```
##              GB_CVC.Phase_2_number_of_times_objects
## median                                       3.000
## mean                                         5.542
## SE.mean                                      1.871
## CI.mean.0.95                                 3.870
## var                                         83.998
## std.dev                                      9.165
## coef.var                                     1.654
## skewness                                     3.283
## skew.2SE                                     3.476
## kurtosis                                    10.368
## kurt.2SE                                     5.649
## normtest.W                                   0.454
## normtest.p                                   0.000
##              GB_GAVC.Phase_2_number_of_times_objects
## median                                        4.000
## mean                                          3.583
## SE.mean                                       0.275
## CI.mean.0.95                                  0.569
## var                                           1.819
```

```
## std.dev                                    1.349
## coef.var                                   0.376
## skewness                                   0.034
## skew.2SE                                   0.036
## kurtosis                                  -0.971
## kurt.2SE                                  -0.529
## normtest.W                                 0.943
## normtest.p                                 0.186
```

The skew.2SE and kurt.2SE for CVC are higher than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis for CVC are significant (at $p < 0.05$). The skew.2SE and kurt.2SE for GAVC are lower than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis for GAVC are not significant (at $p < 0.05$). The p-value (indicated by normtest.p) obtained by the Shapiro-Wilk test for CVC is 0 and thus the CVC is not normally distributed. The p-value (indicated by normtest.p) obtained by the Shapiro-Wilk test for GAVC is >0.05 and thus the GAVC is normally distributed.

Paired difference test

Phase 1: total gaze duration

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the data is normally distributed and thus a paired t-test was used:

```
t.test(GB_CVC$Phase_1_total_time, GB_GAVC$Phase_1_total_time, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  GB_CVC$Phase_1_total_time and GB_GAVC$Phase_1_total_time
## t = -0.26241, df = 23, p-value = 0.7953
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.811726  3.728392
## sample estimates:
## mean of the differences
##              -0.5416667
```

The p-value is 0.80 which is bigger that 0.05, so we can't reject the null hypothesis.

Phase 1: number of gazes

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the data is normally distributed and thus a paired t-test was used:

```
t.test(GB_CVC$Phase_1_number_of_times, GB_GAVC$Phase_1_number_of_times, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  GB_CVC$Phase_1_number_of_times and GB_GAVC$Phase_1_number_of_times
## t = 1.2746, df = 23, p-value = 0.2152
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.219889  5.136556
## sample estimates:
## mean of the differences
```

```
##                1.958333
```

The p-value is 0.22 which is bigger that 0.05, so we can't reject the null hypothesis.

Phase 2: total gaze duration

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the data is normally distributed and thus a paired t-test was used:

```
t.test(GB_CVC$Phase_2_total_time, GB_GAVC$Phase_2_total_time, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  GB_CVC$Phase_2_total_time and GB_GAVC$Phase_2_total_time
## t = 2.3726, df = 23, p-value = 0.02641
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4056473 5.9276861
## sample estimates:
## mean of the differences
##                3.166667
```

The p-value is 0.03 which is smaller than 0.05, so we can reject the null hypothesis. Comparing the mean and median of the Dummy and Special shows that the Dummy has a significantly higher value for phase 2 total time.

Phase 2: number of gazes

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the data is normally distributed and thus a paired t-test was used:

```
t.test(GB_CVC$Phase_2_number_of_times, GB_GAVC$Phase_2_number_of_times, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  GB_CVC$Phase_2_number_of_times and GB_GAVC$Phase_2_number_of_times
## t = 1.6409, df = 23, p-value = 0.1144
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5213086  4.5213086
## sample estimates:
## mean of the differences
##                       2
```

The p-value is 0.11 which is higher than 0.05, so we can't reject the null hypothesis.

Phase 2: number of gazes followed to objects

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the GAVC data is not normally distributed and thus a Wilcoxon signed-rank test was used:

```
wilcox_result <- wilcox.test(GB_CVC$Phase_2_number_of_times_objects, GB_GAVC$Phase_2_number_of_times_ob
            paired = TRUE, exact = TRUE, correct = FALSE)
```

```
## Warning in wilcox.test.default(GB_CVC$Phase_2_number_of_times_objects,
```

```
## GB_GAVC$Phase_2_number_of_times_objects, : cannot compute exact p-value
## with ties
```

```
## Warning in wilcox.test.default(GB_CVC$Phase_2_number_of_times_objects,
## GB_GAVC$Phase_2_number_of_times_objects, : cannot compute exact p-value
## with zeroes
```

```
wilcox_result
```

```
##
##  Wilcoxon signed rank test
##
## data:  GB_CVC$Phase_2_number_of_times_objects and GB_GAVC$Phase_2_number_of_times_objects
## V = 100.5, p-value = 0.5955
## alternative hypothesis: true location shift is not equal to 0
```

The p-value is 0.60 which is higher than 0.05, so we can't reject the null hypothesis.

```
Zstat <- qnorm(wilcox_result$p.value/2)
Zstat
```

```
## [1] -0.530828
```

The Z-value is -0.53, which will be used to calculate the effect size.

Effect size

Phase 1: total gaze duration

The effect size for a parametric test is calculated using Cohen's d:

```
cohensD(GB_CVC$Phase_1_total_time, GB_GAVC$Phase_1_total_time)
```

```
## [1] 0.04854197
```

The effect size (d) is 0.05.

Phase 1: number of gazes

The effect size for a parametric test is calculated using Cohen's d:

```
cohensD(GB_CVC$Phase_1_number_of_times, GB_GAVC$Phase_1_number_of_times)
```

```
## [1] 0.1971675
```

The effect size (d) is 0.20.

Phase 2: total gaze duration

The effect size for a parametric test is calculated using Cohen's d:

```
cohensD(GB_CVC$Phase_2_total_time, GB_GAVC$Phase_2_total_time)
```

```
## [1] 0.3524396
```

The effect size (d) is 0.35.

Phase 2: number of gazes

The effect size for a parametric test is calculated using Cohen's d:

```r
cohensD(GB_CVC$Phase_2_number_of_times, GB_GAVC$Phase_2_number_of_times)
```

```
## [1] 0.2483222
```

The effect size (d) is 0.25.

Phase 2: number of gazes followed to objects

The effect size is calculated using Pearson's r for a non-parametric test:

```r
abs(Zstat)/sqrt(24)
```

```
## [1] 0.1083548
```

The effect size r is 0.11.