What is p-value and Confidence Interval?

Kishore Puthezhath

2022-01-2

Assume that we have a normally distributed population of males and females with a mean height of 168cms and 158cms respectively.

```
library(ggplot2)
set.seed(1)
m=rnorm(100,168,5)
f=rnorm(100,158,5)
mf=as.data.frame(m,f)
```

```
## Warning in as.data.frame.numeric(m, f): 'row.names' is not a character vector of
## length 100 -- omitting it. Will be an error!
```

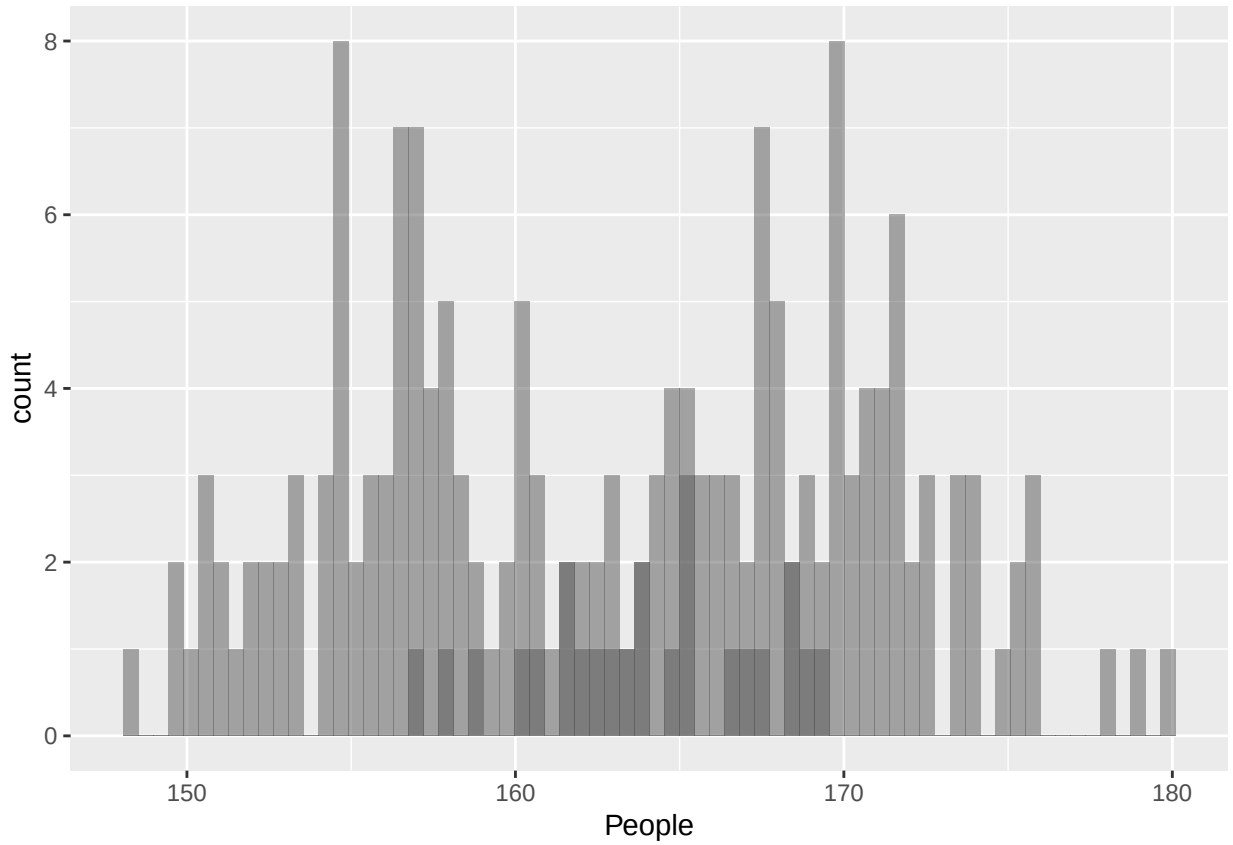
```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.5     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.0.2     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
mf %>%
ggplot()+
  geom_histogram(mapping = aes(m
              ),position = 'identity',bins=70, alpha=.5, )+
  geom_histogram(mapping = aes(f
              ),position = 'identity',bins = 70, alpha=.5)+xlab("People")
```




```
htdif=mean(m)-mean(f)
```

We can see that the mean difference in the height is 10 cms.