STA299

Grace Yu

2022-04-14

```
library(tidyverse)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

10%

```
validation <- read_csv("./validation/10%.csv")
```

```
## Rows: 13 Columns: 32
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (30): All, Femur, Patella, QT, bbox/AP50, bbox/AP75, bbox/APl, bbox/APm,...
## lgl  (2): bbox/APs, bbox_student/APs
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(validation)
```

```
## Rows: 13
## Columns: 32
## $ All                                        <dbl> 16.30732, 18.55672, 18.660~
## $ Femur                                      <dbl> 13.78256, 15.03444, 16.159~
## $ Patella                                    <dbl> 18.10062, 19.67644, 20.873~
## $ QT                                         <dbl> 17.03877, 20.95929, 18.949~
## $ `bbox/AP50`                                <dbl> 53.83413, 62.92059, 62.185~
## $ `bbox/AP75`                                <dbl> 5.683726, 4.534259, 5.8276~
## $ `bbox/APl`                                 <dbl> 21.00806, 22.28706, 26.371~
## $ `bbox/APm`                                 <dbl> 14.82817, 16.63017, 14.166~
## $ `bbox/APs`                                 <lgl> NA, NA, NA, NA, NA, NA, NA~
## $ `bbox_student/AP`                          <dbl> 15.90876, 17.02320, 17.327~
## $ `bbox_student/AP-Femur`                    <dbl> 16.09805, 14.15410, 13.370~
## $ `bbox_student/AP-Patella`                  <dbl> 12.10712, 18.49853, 17.259~
## $ `bbox_student/AP-Quadriceps Tendon`        <dbl> 19.52111, 18.41695, 21.353~
## $ `bbox_student/AP50`                        <dbl> 54.84487, 55.55384, 59.580~
## $ `bbox_student/AP75`                        <dbl> 3.283055, 3.179555, 2.8548~
## $ `bbox_student/APl`                         <dbl> 21.73136, 22.45724, 24.526~
## $ `bbox_student/APm`                         <dbl> 10.921054, 11.750413, 9.68~
```

```
## $ `bbox_student/APs`                              <lgl> NA, NA, NA, NA, NA, NA, NA~
## $ data_time                                       <dbl> 0.03137388, 0.03561907, 0.~
## $ eta_seconds                                     <dbl> 465718.7, 459761.7, 458085~
## $ iteration                                       <dbl> 2999, 3999, 4999, 5999, 69~
## $ loss_box_reg                                    <dbl> 0.13882254, 0.12651279, 0.~
## $ loss_cls                                        <dbl> 0.03848956, 0.03578251, 0.~
## $ loss_rpn_cls                                    <dbl> 0.04239071, 0.03908001, 0.~
## $ loss_rpn_loc                                    <dbl> 0.01387088, 0.01331944, 0.~
## $ lr                                              <dbl> 0.014110219, 0.013033299, ~
## $ `roi_head/num_target_bg_samples_supervised`     <dbl> 0.008733911, 0.007542802, ~
## $ `roi_head/num_target_fg_samples_supervised`     <dbl> 0.04645591, 0.05023066, 0.~
## $ `rpn/num_neg_anchors`                           <dbl> 0.01245394, 0.01118025, 0.~
## $ `rpn/num_pos_anchors`                           <dbl> 0.01, 0.01, 0.01, 0.01, 0.~
## $ time                                            <dbl> 494.9375, 494.8125, 486.56~
## $ total_loss                                      <dbl> 17.0625, 17.1875, 25.4375,~
```

```r
# Summary table
validation %>%
  dplyr::select(All, Femur, Patella, QT)%>%
  summarize(n=n(),
            value = c("mean", "median"),
            AP = c(mean(All), median(All)),
            Femur = c(mean(Femur), median(Femur)),
            Patella = c(mean(Patella), median(Patella)),
            QT = c(mean(QT), median(QT)))
```

```
## # A tibble: 2 x 6
##       n value     AP Femur Patella    QT
##   <int> <chr>  <dbl> <dbl>   <dbl> <dbl>
## 1    13 mean    20.0  16.0    22.9  21.1
## 2    13 median  20.5  15.9    23.9  21.6
```

Use R to find a 95% bootstrap confidence interval for the *mean*. Use 1000 bootstrap samples. *NOTE:* More bootstrap samples is better, but if you find this times out or takes too long in RStudio Cloud, try using 1000 bootstrap samples instead.

```r
repetitions <- 1000
All_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Femur_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Patella_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
QT_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
sample_size <- as.numeric(validation %>% summarize(n()))
set.seed(50)
for (i in 1:repetitions)
{
  boot_samp <- validation %>% sample_n(size = sample_size, replace=TRUE)
  All_means[i] <- as.numeric(boot_samp %>% summarize(mean(All)))
  Femur_means[i] <- as.numeric(boot_samp %>% summarize(mean(Femur)))
  Patella_means[i] <- as.numeric(boot_samp %>% summarize(mean(Patella)))
  QT_means[i] <- as.numeric(boot_samp %>% summarize(mean(QT)))
}
quantile(All_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 19.20709 20.62560
```

```
quantile(Femur_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 15.31994 16.85163
```

```
quantile(Patella_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 21.57187 24.04805
```

```
quantile(QT_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 20.16798 21.81854
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

5%

```
validation <- read_csv("./validation/5%.csv")
```

```
## Rows: 13 Columns: 4
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (4): All, Femur, Patella, QT
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(validation)
```

```
## Rows: 13
## Columns: 4
## $ All     <dbl> 15.34879, 16.10447, 16.72992, 17.74713, 18.05024, 18.18740, 17~
## $ Femur   <dbl> 13.73200, 13.81519, 14.31970, 14.84224, 14.91372, 14.99309, 14~
## $ Patella <dbl> 15.26605, 16.32484, 16.96122, 19.01103, 20.20069, 20.09757, 20~
## $ QT      <dbl> 17.04833, 18.17339, 18.90885, 19.38811, 19.03630, 19.47154, 18~
```

```
# Summary table
validation %>%
  dplyr::select(All, Femur, Patella, QT)%>%
  summarize(n=n(),
            value = c("mean", "median"),
            AP = c(mean(All), median(All)),
            Femur = c(mean(Femur), median(Femur)),
            Patella = c(mean(Patella), median(Patella)),
            QT = c(mean(QT), median(QT)))
```

```
## # A tibble: 2 x 6
##       n value     AP Femur Patella    QT
##   <int> <chr>  <dbl> <dbl>   <dbl> <dbl>
## 1    13 mean    17.0  13.9    18.8  18.5
## 2    13 median  17.1  14.3    19.0  18.8
```

Use R to find a 95% bootstrap confidence interval for the *mean*. Use 1000 bootstrap samples. *NOTE:* More bootstrap samples is better, but if you find this times out or takes too long in RStudio Cloud, try using 1000 bootstrap samples instead.

```
repetitions <- 1000
All_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Femur_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Patella_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
QT_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
sample_size <- as.numeric(validation %>% summarize(n()))
set.seed(50)
for (i in 1:repetitions)
{
  boot_samp <- validation %>% sample_n(size = sample_size, replace=TRUE)
  All_means[i] <- as.numeric(boot_samp %>% summarize(mean(All)))
  Femur_means[i] <- as.numeric(boot_samp %>% summarize(mean(Femur)))
  Patella_means[i] <- as.numeric(boot_samp %>% summarize(mean(Patella)))
  QT_means[i] <- as.numeric(boot_samp %>% summarize(mean(QT)))
}
quantile(All_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 16.48219 17.55811
```

```
quantile(Femur_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 13.36616 14.43350
```

```
quantile(Patella_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 17.82256 19.67808
```

```
quantile(QT_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 17.89430 18.90589
```

2%

```
validation <- read_csv("./validation/2%.csv")
```

```
## Rows: 13 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (5): All, Femur, Patella, QT, iteration
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(validation)
```

```
## Rows: 13
## Columns: 5
## $ All       <dbl> 12.07271, 12.03902, 11.98398, 12.53700, 12.23104, 12.21089, ~
## $ Femur     <dbl> 10.69432, 11.11448, 11.58567, 11.36833, 10.50030, 11.35741, ~
## $ Patella   <dbl> 13.07103, 12.50941, 11.68044, 13.03049, 12.22614, 12.15484, ~
## $ QT        <dbl> 12.45278, 12.49317, 12.68582, 13.21219, 13.96669, 13.12043, ~
## $ iteration <dbl> 2999, 3999, 4999, 5999, 6999, 7999, 8999, 9999, 10999, 11999~
```

```r
# Summary table
validation %>%
  dplyr::select(All, Femur, Patella, QT)%>%
  summarize(n=n(),
            value = c("mean", "median"),
            AP = c(mean(All), median(All)),
            Femur = c(mean(Femur), median(Femur)),
            Patella = c(mean(Patella), median(Patella)),
            QT = c(mean(QT), median(QT)))
```

```
## # A tibble: 2 x 6
##       n value     AP Femur Patella    QT
##   <int> <chr>  <dbl> <dbl>   <dbl> <dbl>
## 1    13 mean    12.2  11.2    11.8  13.5
## 2    13 median  12.1  11.3    11.7  13.4
```

Use R to find a 95% bootstrap confidence interval for the *mean*. Use 1000 bootstrap samples. *NOTE:* More bootstrap samples is better, but if you find this times out or takes too long in RStudio Cloud, try using 1000 bootstrap samples instead.

```r
repetitions <- 1000
All_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Femur_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Patella_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
QT_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
sample_size <- as.numeric(validation %>% summarize(n()))
set.seed(50)
for (i in 1:repetitions)
{
  boot_samp <- validation %>% sample_n(size = sample_size, replace=TRUE)
  All_means[i] <- as.numeric(boot_samp %>% summarize(mean(All)))
  Femur_means[i] <- as.numeric(boot_samp %>% summarize(mean(Femur)))
  Patella_means[i] <- as.numeric(boot_samp %>% summarize(mean(Patella)))
  QT_means[i] <- as.numeric(boot_samp %>% summarize(mean(QT)))
}
quantile(All_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 12.05233 12.28151
```

```r
quantile(Femur_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 10.95957 11.36661
```

```r
quantile(Patella_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 11.42101 12.19456
```

```r
quantile(QT_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 13.08295 13.99471
```

1%

```
validation <- read_csv("./validation/1%.csv")
```

```
## Rows: 13 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (5): All, Femur, Patella, QT, iteration
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(validation)
```

```
## Rows: 13
## Columns: 5
## $ All       <dbl> 8.788097, 11.219007, 11.768174, 11.318499, 11.193552, 11.134~
## $ Femur     <dbl> 11.33304, 14.80275, 15.21662, 13.09089, 15.00638, 13.84156, ~
## $ Patella   <dbl> 5.752117, 8.083808, 7.941371, 7.789806, 5.649803, 5.645825, ~
## $ QT        <dbl> 9.279135, 10.770466, 12.146534, 13.074801, 12.924476, 13.916~
## $ iteration <dbl> 2999, 3999, 4999, 5999, 6999, 7999, 8999, 9999, 10999, 11999~
```

```
# Summary table
validation %>%
  dplyr::select(All, Femur, Patella, QT)%>%
  summarize(n=n(),
            value = c("mean", "median"),
            AP = c(mean(All), median(All)),
            Femur = c(mean(Femur), median(Femur)),
            Patella = c(mean(Patella), median(Patella)),
            QT = c(mean(QT), median(QT)))
```

```
## # A tibble: 2 x 6
##       n value     AP Femur Patella    QT
##   <int> <chr>  <dbl> <dbl>   <dbl> <dbl>
## 1    13 mean    11.3  14.3    6.44  13.1
## 2    13 median  11.4  14.4    6.03  13.7
```

Use R to find a 95% bootstrap confidence interval for the *mean*. Use 1000 bootstrap samples. *NOTE:* More bootstrap samples is better, but if you find this times out or takes too long in RStudio Cloud, try using 1000 bootstrap samples instead.

```
repetitions <- 1000
All_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Femur_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
Patella_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
QT_means <- rep(NA, repetitions)  # where we'll store the bootstrap means
sample_size <- as.numeric(validation %>% summarize(n()))
set.seed(50)
for (i in 1:repetitions)
{
  boot_samp <- validation %>% sample_n(size = sample_size, replace=TRUE)
  All_means[i] <- as.numeric(boot_samp %>% summarize(mean(All)))
  Femur_means[i] <- as.numeric(boot_samp %>% summarize(mean(Femur)))
  Patella_means[i] <- as.numeric(boot_samp %>% summarize(mean(Patella)))
  QT_means[i] <- as.numeric(boot_samp %>% summarize(mean(QT)))
}
```

```r
quantile(All_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 10.80489 11.62582
```

```r
quantile(Femur_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 13.54166 14.88333
```

```r
quantile(Patella_means,c(0.025,0.975))
```

```
##    2.5%   97.5%
## 5.95637 7.01500
```

```r
quantile(QT_means,c(0.025,0.975))
```

```
##     2.5%    97.5%
## 12.26082 13.90118
```