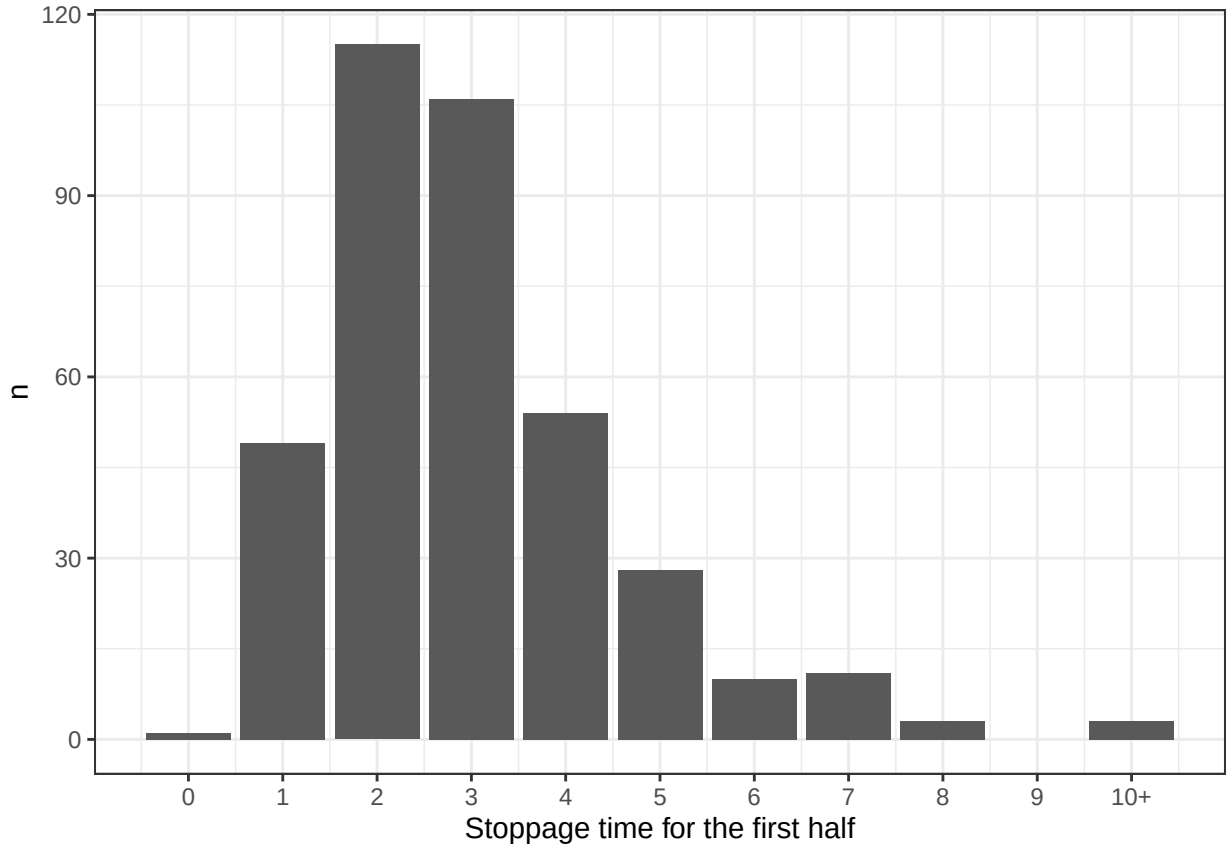
Stoppage time models for the 2020 season

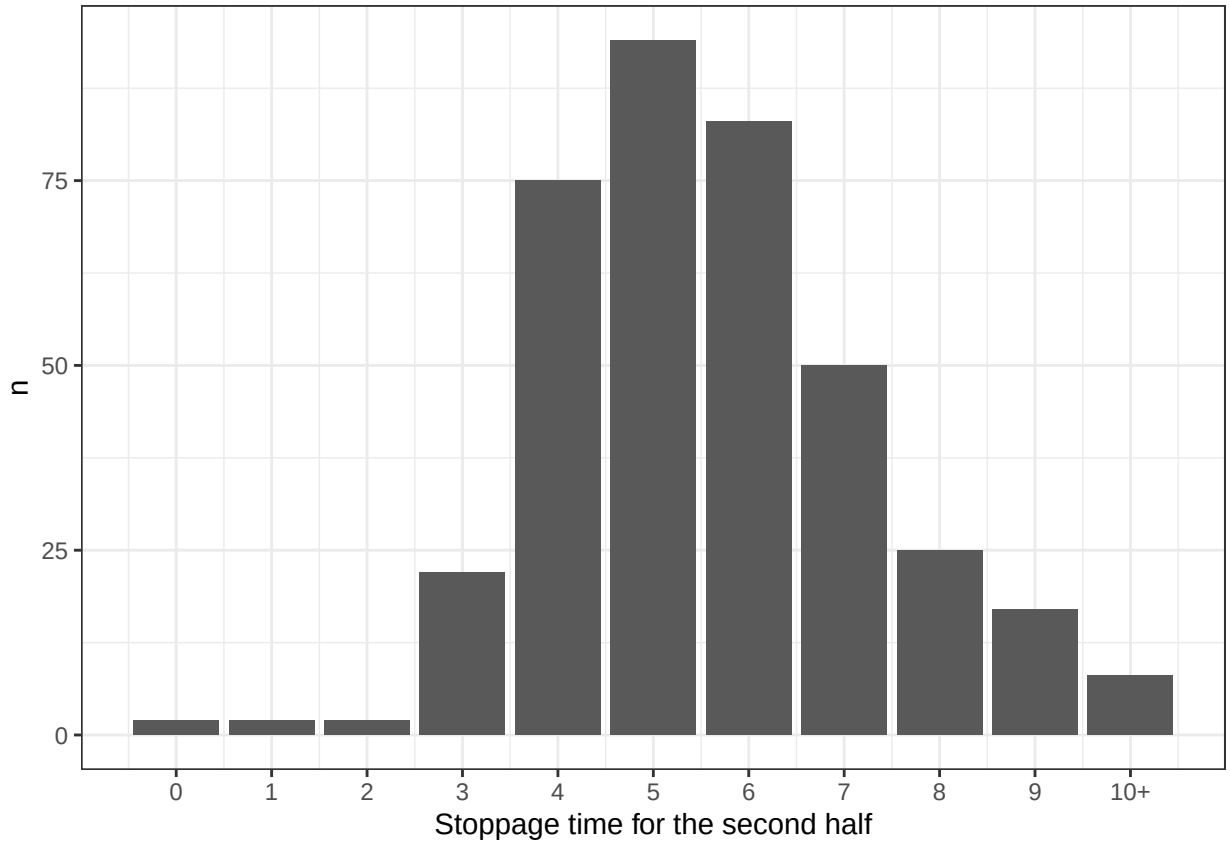
```
library(dplyr)
library(ggplot2)
library(CVXR)

load("2020/data/input.RData")
```

```
st1 = U1
st1[which(st1 > 10)] = 10
tibble(st1) %>%
  count(st1) %>%
  ggplot(aes(x = st1, y = n)) +
  geom_bar(position = "dodge", stat = "identity") +
  theme_bw() +
  xlab("Stoppage time for the first half") +
  ylab("n") +
  scale_x_continuous(breaks = 0:10,
                     labels = c(0:9, "10+"))
```



```r
st2 = U2
st2[which(st2 > 10)] = 10
tibble(st2) %>%
  count(st2) %>%
  ggplot(aes(x = st2, y = n)) +
  geom_bar(position = "dodge", stat = "identity") +
  theme_bw() +
  xlab("Stoppage time for the second half") +
  ylab("n") +
  scale_x_continuous(breaks = 0:10,
                     labels = c(0:9, "10+"))
```



The stoppage time for the first half, U_1, and the second half, U_2, are modeled as:

$$U_1 \sim \text{Poisson}(\pi_1)$$

$$U_2 \sim \text{Poisson}(\pi_2)$$

- g_t is the amount of goals scored in half t until minute 45;
- r_t is the amount of red cards received in half t until minute 45;
- $c = \begin{cases} 1, & \text{if } |x - y| \leq 1 \text{ at minute 45 of the second half;} \\ 0, & \text{otherwise.} \end{cases}$

```r
aic <- function(loglik, k) {
  2*k - 2*loglik
}

bic <- function(loglik, k, n) {
  k*log(n) - 2*loglik
}
```

Model 1

$$\pi_1 = \eta_1$$

$$\pi_2 = \eta_2$$

```r
eta = Variable(2)
pi1 = eta[1] * rep(1, N)
pi2 = eta[2] * rep(1, N)
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_1 = list(eta = as.vector(solution$getValue(eta)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_1$aic = aic(mod_1$loglik, 2)
mod_1$bic = bic(mod_1$loglik, 2, N)
mod_1
```

```
## $eta
## [1] 3.015829 5.597358
##
## $loglik
## [1] -1476.224
##
## $aic
## [1] 2956.448
##
## $bic
## [1] 2964.329
```

Model 2

$$\pi_1 = \eta_1 + \phi g_1$$

$$\pi_2 = \eta_2 + \phi g_2$$

```
eta = Variable(2)
phi = Variable(1)
pi1 = eta[1] + phi * g1
pi2 = eta[2] + phi * g2
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_2 = list(eta = as.vector(solution$getValue(eta)),
             phi = as.vector(solution$getValue(phi)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_2$aic = aic(mod_2$loglik, 3)
mod_2$bic = bic(mod_2$loglik, 3, N)
mod_2
```

```
## $eta
## [1] 3.019638 5.601423
##
## $phi
## [1] -0.003621193
##
## $loglik
## [1] -1476.223
##
## $aic
## [1] 2958.446
##
## $bic
## [1] 2970.266
```

Model 3

$$\pi_1 = \eta_1 + \phi_1 g_1$$
$$\pi_2 = \eta_2 + \phi_2 g_2$$

```
eta = Variable(2)
phi = Variable(2)
pi1 = eta[1] + phi[1] * g1
pi2 = eta[2] + phi[2] * g2
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_3 = list(eta = as.vector(solution$getValue(eta)),
             phi = as.vector(solution$getValue(phi)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_3$aic = aic(mod_3$loglik, 4)
mod_3$bic = bic(mod_3$loglik, 4, N)
mod_3
```

```
## $eta
## [1] 2.958049 5.707686
##
## $phi
## [1]  0.05556236 -0.09681558
##
## $loglik
## [1] -1475.663
##
## $aic
## [1] 2959.325
##
## $bic
## [1] 2975.086
```

Model 4

$$\pi_1 = \eta_1 + \rho r_1$$

$$\pi_2 = \eta_2 + \rho r_2$$

```
eta = Variable(2)
rho = Variable(1)
pi1 = eta[1] + rho * r1
pi2 = eta[2] + rho * r2
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_4 = list(eta = as.vector(solution$getValue(eta)),
             rho = as.vector(solution$getValue(rho)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_4$aic = aic(mod_4$loglik, 3)
mod_4$bic = bic(mod_4$loglik, 3, N)
mod_4
```

```
## $eta
## [1] 2.965075 5.502409
##
## $rho
## [1] 0.779728
##
## $loglik
## [1] -1471.209
##
## $aic
## [1] 2948.419
##
## $bic
## [1] 2960.239
```

Model 5

$$\pi_1 = \eta_1 + \rho_1 r_1$$

$$\pi_2 = \eta_2 + \rho_2 r_2$$

```
eta = Variable(2)
rho = Variable(2)
pi1 = eta[1] + rho[1] * r1
pi2 = eta[2] + rho[2] * r2
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_5 = list(eta = as.vector(solution$getValue(eta)),
             rho = as.vector(solution$getValue(rho)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_5$aic = aic(mod_5$loglik, 4)
mod_5$bic = bic(mod_5$loglik, 4, N)
mod_5
```

```
## $eta
## [1] 2.922251 5.580481
##
## $rho
## [1] 1.8708853 0.1234095
##
## $loglik
## [1] -1466.406
##
## $aic
## [1] 2940.812
##
## $bic
## [1] 2956.573
```

Model 6

$$\pi_1 = \eta_1 + \rho_1 r_1$$

$$\pi_2 = \eta_2 + \rho_2 r_2 + \kappa c$$

```
eta = Variable(2)
rho = Variable(2)
kappa = Variable(1)
pi1 = eta[1] + rho[1] * r1
pi2 = eta[2] + rho[2] * r2 + kappa * c
log_lik_st = t(U1) %*% log(pi1) + t(U2) %*% log(pi2) - sum_entries(pi1) - sum_entries(pi2)
objective = Maximize(log_lik_st)
problem = Problem(objective)
solution = solve(problem, solver = "MOSEK")
mod_6 = list(eta = as.vector(solution$getValue(eta)),
             rho = as.vector(solution$getValue(rho)),
             kappa = as.vector(solution$getValue(kappa)),
             loglik = solution$value - sum(log(factorial(U1))) - sum(log(factorial(U2))))
mod_6$aic = aic(mod_6$loglik, 5)
mod_6$bic = bic(mod_6$loglik, 5, N)
mod_6
```

```
## $eta
## [1] 2.922272 4.735503
##
## $rho
## [1] 1.8708638 0.1345932
##
## $kappa
## [1] 1.18707
##
## $loglik
## [1] -1456.245
##
## $aic
## [1] 2922.49
##
## $bic
## [1] 2942.191
```

```
sum(g1)
```

```
## [1] 395
```

```
sum(g2)
```

```
## [1] 433
```

```
sum(r1)
```

```
## [1] 19
```

```
sum(r2)
```

```
## [1] 52
```

```
sum(c)
```

```
## [1] 270
```