Plotting HR Frequencies by Park

Jason Osborne

2023-06-08

Some preliminary data processing is not covered here, but we begin by loading a (60×4) data frame of averages over the 2012-2021 time period whose first several rows are shown below:

```
load("allhrtables.RData")
allhrtables %>% head
```

```
##     hr BAT_HAND_CD park       Freq     fitted season
## 1 TRUE           L  ANA 0.02255985 0.01135641   2012
## 2 TRUE           R  ANA 0.02990654 0.02352751   2012
## 3 TRUE           L  ARI 0.02572479 0.02507295   2012
## 4 TRUE           R  ARI 0.03006501 0.03384510   2012
## 5 TRUE           L  ATL 0.02653343 0.02755824   2012
## 6 TRUE           R  ATL 0.01896933 0.03159598   2012
```

Then we compute averages over the 10 year time period from 2012-2021:

```
allhrtables %>% group_by(park, BAT_HAND_CD) %>%
  summarize(Freq_mean=mean(Freq),fitted_mean=mean(fitted)) ->
  allhrtables.avgs
```

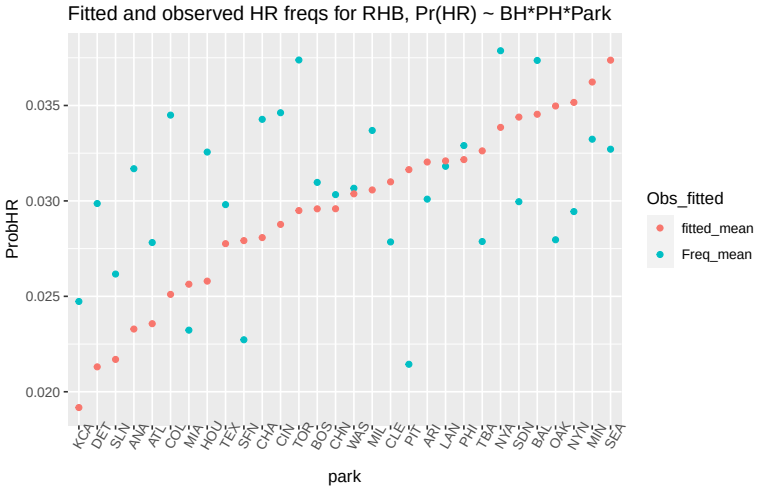
```
## `summarise()` has grouped output by 'park'. You can override using the
## `.groups` argument.
```

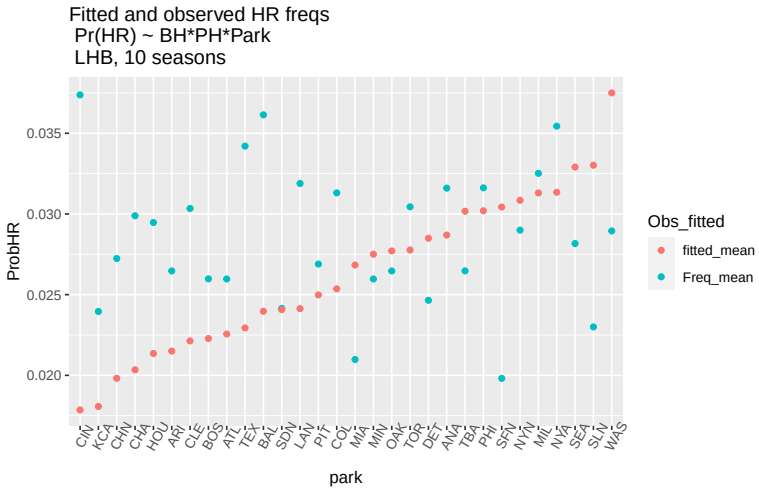
Now, we will reproduce the plot of observed and fitted home run frequencies against park, separately for RHB and LHB.

```
allhrtables.avgs %>% filter(BAT_HAND_CD=="R") -> allhrtables.avgs.R
allhrtables.avgs %>% filter(BAT_HAND_CD=="L") -> allhrtables.avgs.L

allhrtables.avgs.R$park <- factor(allhrtables.avgs.R$park,
                                  levels=allhrtables.avgs.R$park[
                                    order(allhrtables.avgs.R$fitted_mean)])

allhrtables.avgs.R %>%
  pivot_longer(cols=c(Freq_mean,fitted_mean),values_to="ProbHR",names_to="Obs_fitted") ->
allhrtables.avgs.R.wide

allhrtables.avgs.L$park <- factor(allhrtables.avgs.L$park,
                                  levels=allhrtables.avgs.L$park[
                                    order(allhrtables.avgs.L$fitted_mean)])

allhrtables.avgs.L %>%
  pivot_longer(cols=c(Freq_mean,fitted_mean),values_to="ProbHR",names_to="Obs_fitted") ->
allhrtables.avgs.L.wide
```

Not we'll ggplot them:

```
ggplot(allhrtables.avgs.R.wide,aes(y=ProbHR,x=park,color=Obs_fitted)) +
  geom_point() +
  ggtitle("Fitted and observed HR freqs for RHB, Pr(HR) ~ BH*PH*Park") +
  theme(axis.text.x=element_text(angle=60))
```






```
library(knitr)
knit_exit()
```

Now, I'll repeat the construction from last week that averages over batting hand:

```
allhrtables %>% group_by(park) %>%
  summarize(Freq_mean=mean(Freq),fitted_mean=mean(fitted)) ->
  allhrtables.bothavgs

allhrtables.bothavgs$park <- factor(allhrtables.bothavgs$park,
  levels=allhrtables.bothavgs$park[
    order(allhrtables.bothavgs$fitted_mean)])

allhrtables.bothavgs %>%
  pivot_longer(cols=c(Freq_mean,fitted_mean),values_to="ProbHR",names_to="Obs_fitted") ->
allhrtables.bothavgs.wide
```

```
ggplot(allhrtables.bothavgs.wide,aes(y=ProbHR,x=park,color=Obs_fitted)) +
  geom_point() +
  ggtitle("Fitted and observed HR freqs \n Pr(HR) ~ BH*PH*Park \n RHB, 10 seasons") +
  theme(axis.text.x=element_text(angle=60))
```