WU #10 - SEs

Math 154 - Jo Hardin

Tuesday, October 5, 2021

Name: _____

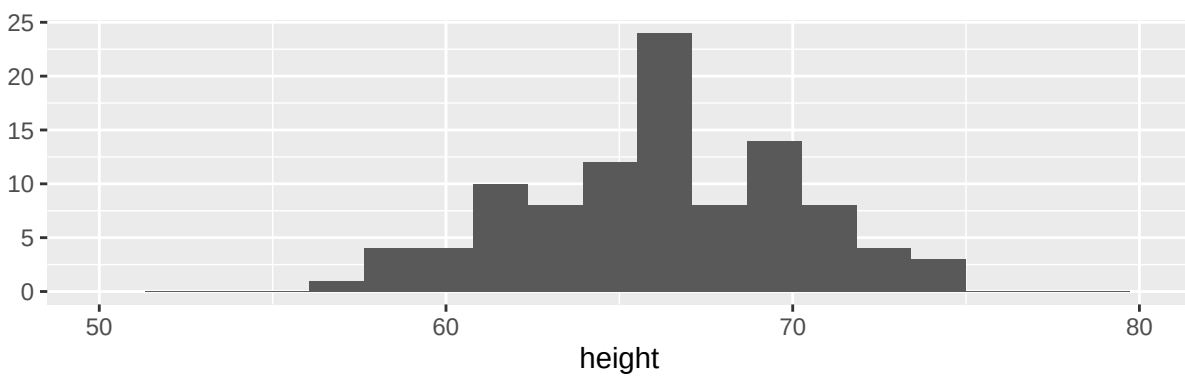
Consider a situation where you are trying to learn about the heights of Claremont Colleges students. The parameter of interest is the true median height of all undergraduates. You'll use the median as the statistic on the observed data. Sketch the following 3 plots (as densities / histograms):

1. The distribution of the heights in a sample of size 100. In inches, mark the center of the curve as well as some indication of the spread of the heights.
2. As a class project, you and your 50 best friends each collect samples of size 100. Sketch the distribution of the 50 sample medians. In inches, mark the center of the curve as well as some indication of the spread of the sample medians.
3. You all run out of time, and end up sampling only 10 undergraduates each. Sketch the distribution of the 50 sample medians. In inches, mark the center of the curve as well as some indication of the spread of the sample medians.

Solution

1.

```
set.seed(47)
data.frame(height = rnorm(100, 66, 4)) %>%
  ggplot() +
  geom_histogram(aes(x = height), bins = 20) +
  ylab("") + xlim(c(50,80))
```



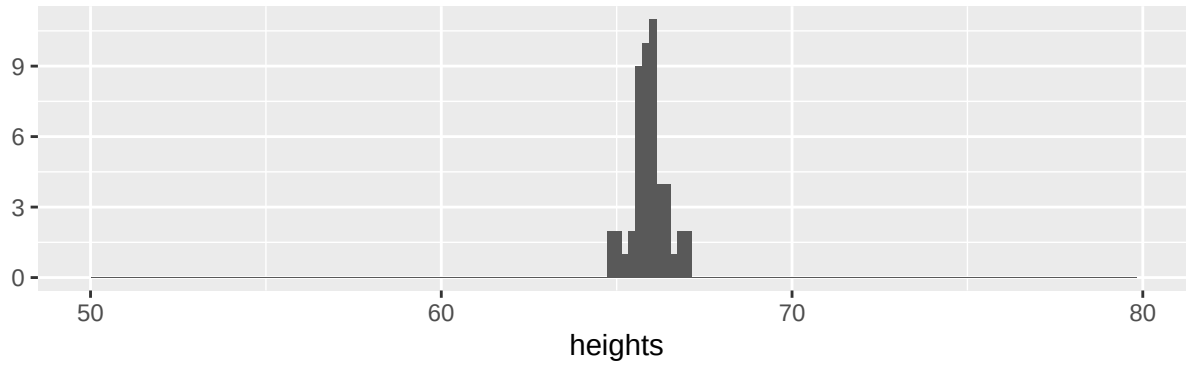
2.

```
set.seed(47)
reps <- 50

norm_func <- function(n_obs){
  data.frame(heights = median(rnorm(n_obs,66,4)) )
```

```
  }

map_df(rep(100, reps), norm_func) %>%
  ggplot() +
  geom_histogram(aes(x = heights), bins = 150) +
  ylab("") + xlim(c(50,80))
```



3.

```
set.seed(10)
reps <- 50
map_df(rep(10, reps), norm_func) %>%
  ggplot() +
  geom_histogram(aes(x = heights), bins = 20) +
  ylab("") + xlim(c(50,80))
```