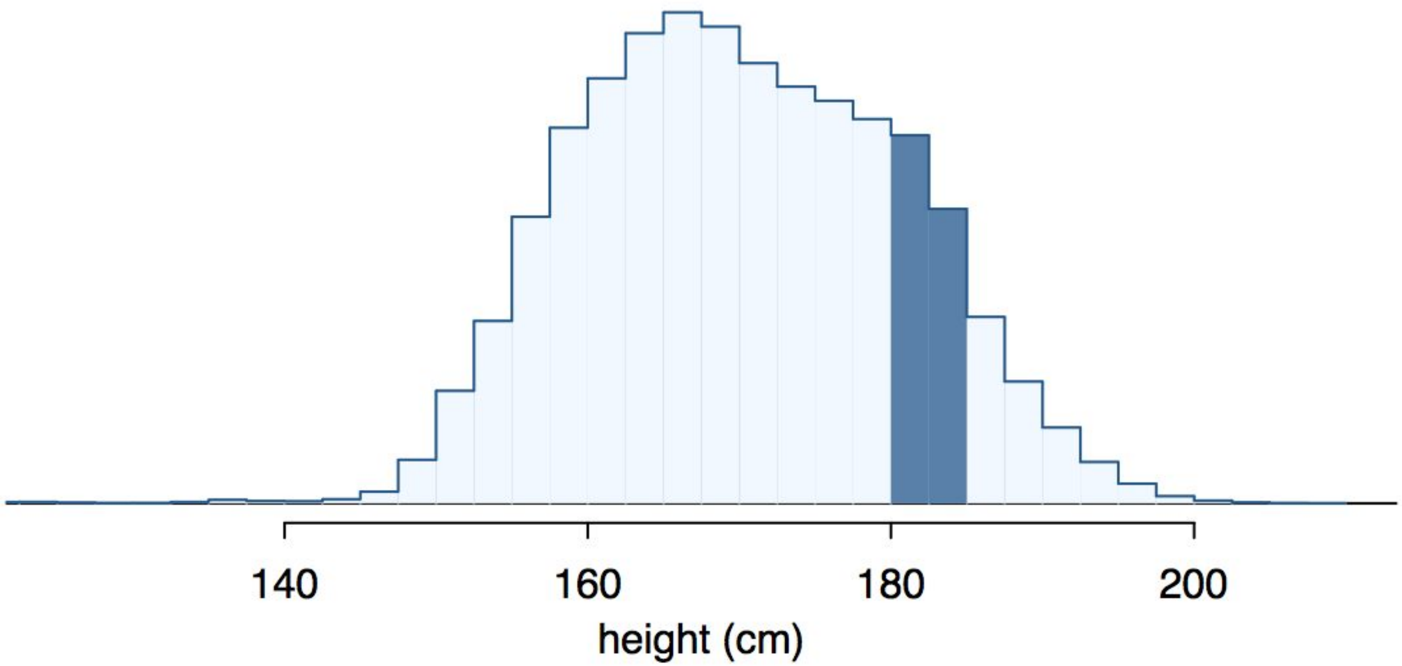
Section 3.5
Continuous Distributions

Stats 7 Summer Session II 2022

Continuous distributions

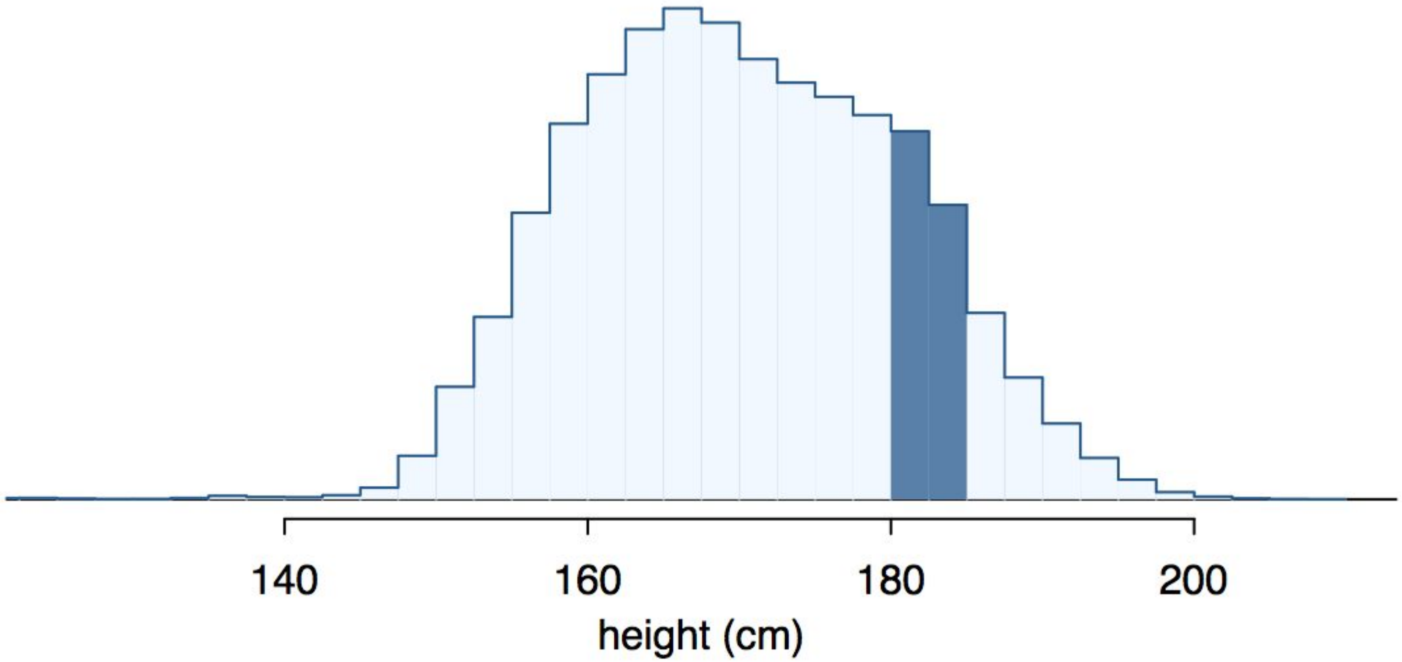
- Below is a histogram of the distribution of sampled heights of US adults
- We can calculate the proportion of data that fall in a given interval from the counts in the bins and the overall total count



Continuous distributions

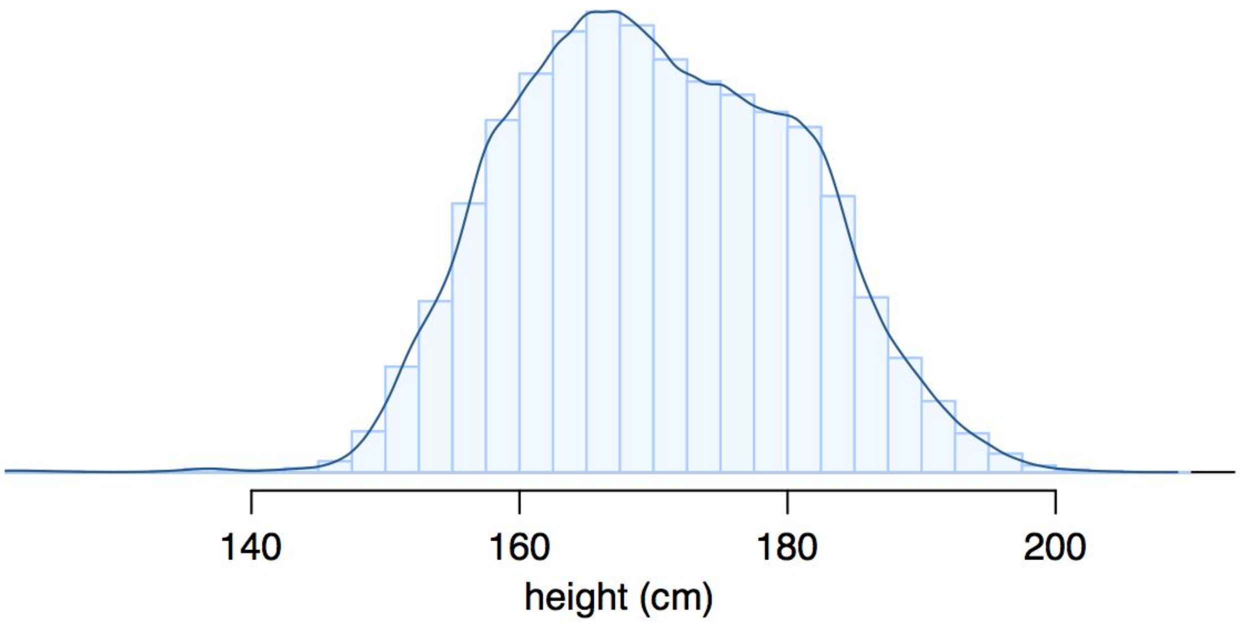
Consider the proportion of adults between 180 cm and 185 cm (about 5'11" to 6'1")

- The two shaded bins have counts of 195,307 and 156,239 people, with a total of 3 million people measured
- (195307 + 156239) / 3000000 = 0.1172
- This gives an estimate of the probability that a randomly sampled US adult is between 180 cm and 185 cm



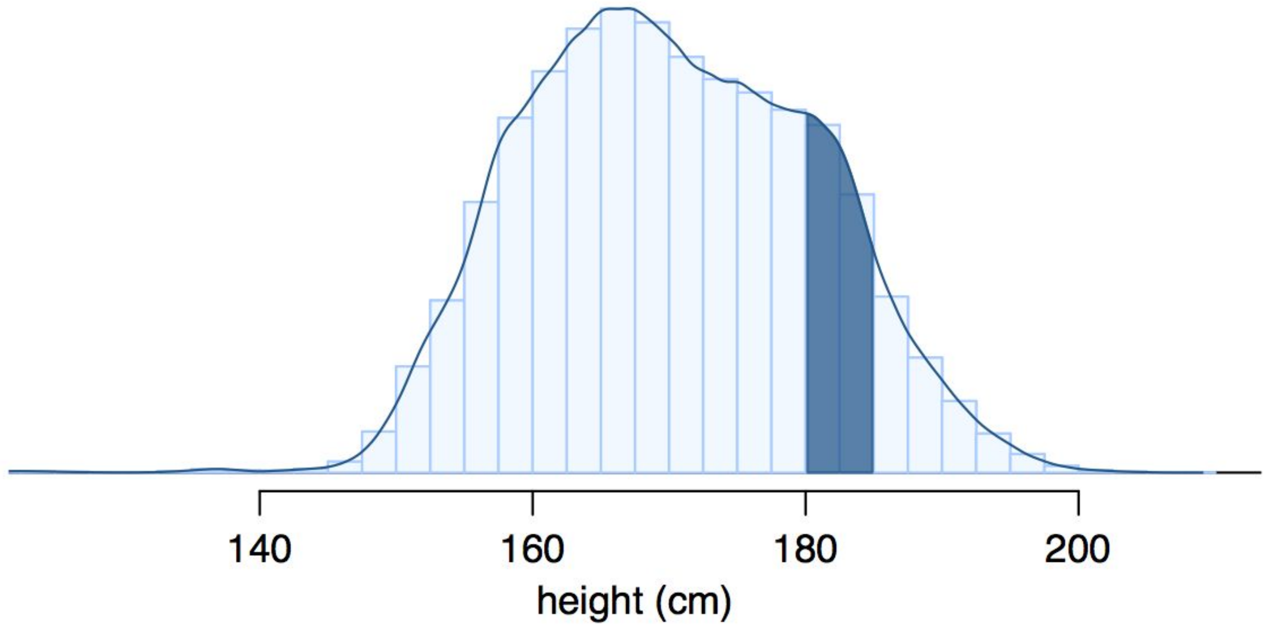
From histograms
to continuous distributions

- This smooth curve represents a *probability density function* (*pdf*), also called a *density* or *distribution*
- A special property of a pdf is the total area under the curve is 1



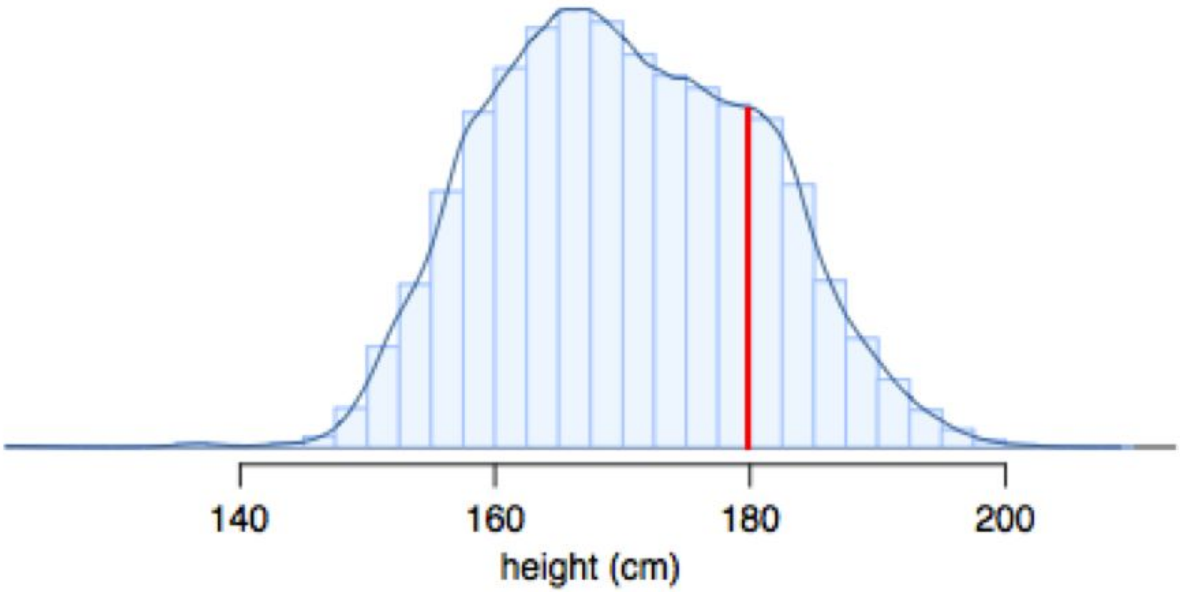
Probabilities from continuous distributions

Therefore, the probability that a randomly sampled US adult is between 180 cm and 185 cm can also be estimated as the shaded area under the curve.



By definition...

Since continuous probabilities are estimated as “the area under the curve”, the probability of a person being exactly 180 cm (or any exact value) is defined as 0.



Why we care

- There are a lot of known probability density functions were we can easily compute the probability of an observation falling in a specified region
- If we believe the distribution of a variable can be approximated by a known pdf, we can use that known pdf as a model to obtain probabilities
- Later, we will see that statistical tests are developed for certain probability densities