Bootstrap Confidence Intervals

Jialiang Mao

February 8, 2016

Overview

- What and why?
- How?
- Always good?

Bootstrapping

- The saying "to pull oneself up by one's bootstraps" was already in use during the 19th century as an example of an impossible task.

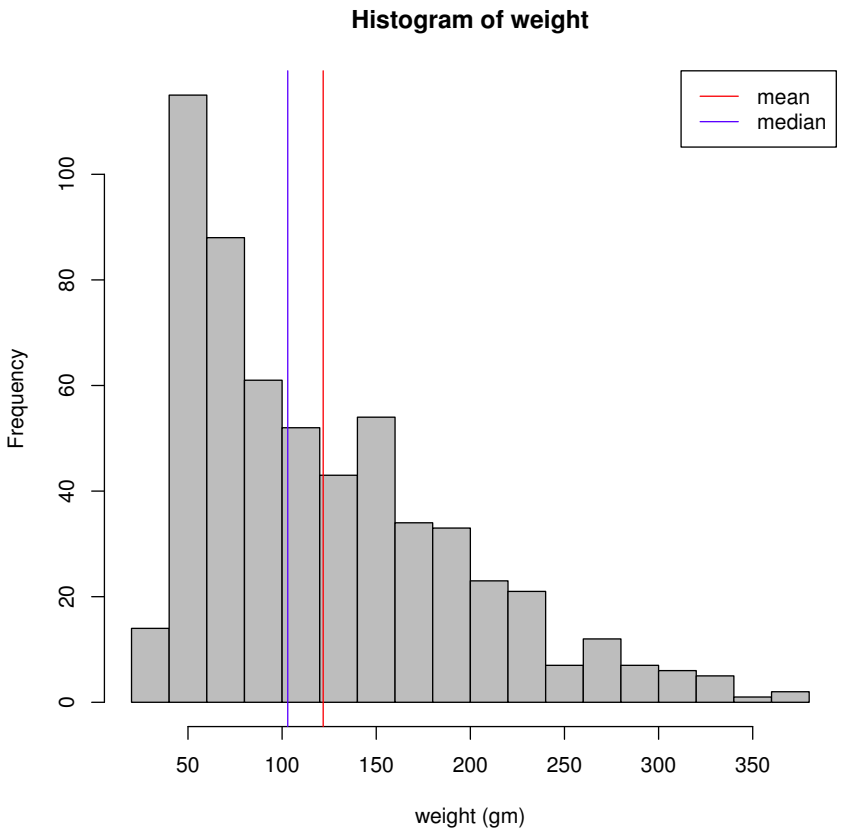
Bootstrapping

- The saying "to pull oneself up by one's bootstraps" was already in use during the 19th century as an example of an impossible task.
- Bootstrap as a metaphor, meaning to better oneself by one's own unaided efforts, was in use in 1922.

Example: Chick weight

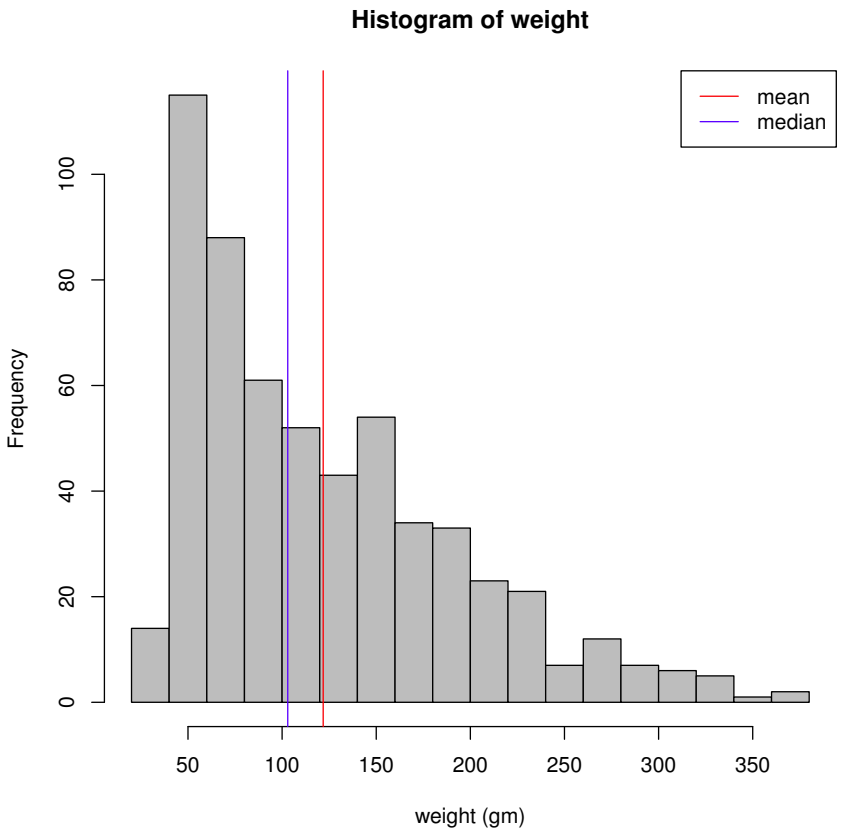
- 578 observations

Example: Chick weight



- 578 observations
- Right-skewed

Example: Chick weight



- 578 observations
- Right-skewed
- median=103

Bootstrap samples

Suppose we are given a sample $(x_1, \ldots, x_n)$ and we want to construct a bootstrap confidence interval for a statistic T (mean, median,...).

Bootstrap samples

Suppose we are given a sample $(x_1, \ldots, x_n)$ and we want to construct a bootstrap confidence interval for a statistic T (mean, median,...).

- a random sample;

Bootstrap samples

Suppose we are given a sample $(x_1, \ldots, x_n)$ and we want to construct a bootstrap confidence interval for a statistic T (mean, median,...).

- a random sample;
- same size;

Bootstrap samples

Suppose we are given a sample $(x_1, \ldots, x_n)$ and we want to construct a bootstrap confidence interval for a statistic T (mean, median,...).

- a random sample;
- same size;
- sample with replacement.

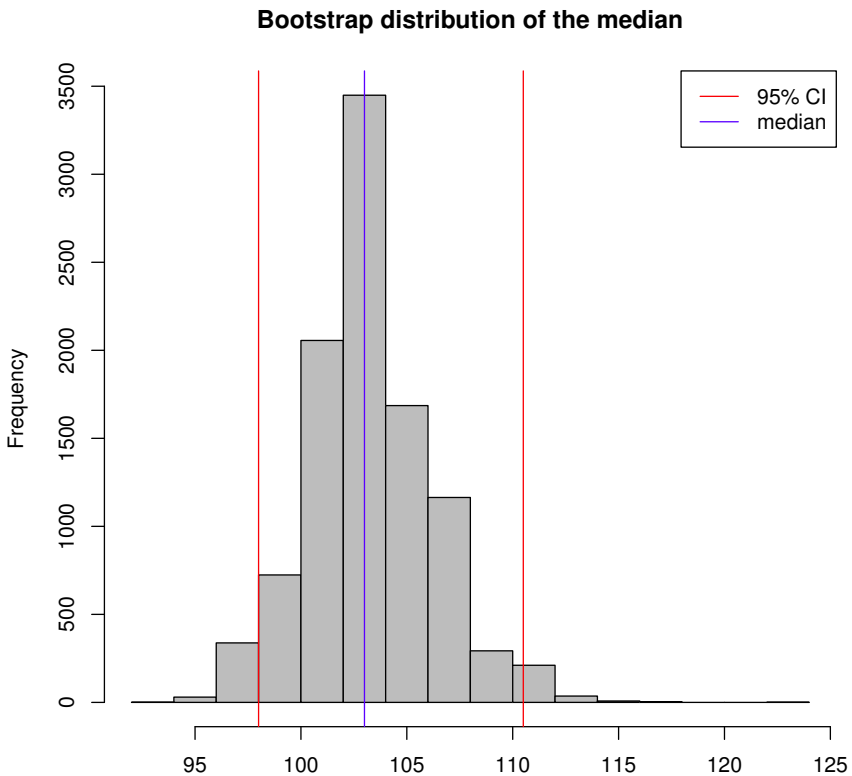
Construct a Bootstrap confidence interval

Take a large number N (for example, $N = 10000$). For $1 \leq i \leq N$,

1. Get a bootstrap sample $(x_1^{*(i)}, \ldots, x_n^{*(i)})$;
2. Calculate the bootstrap statistic $T^{*(i)}$ with the bootstrap sample in the first step.

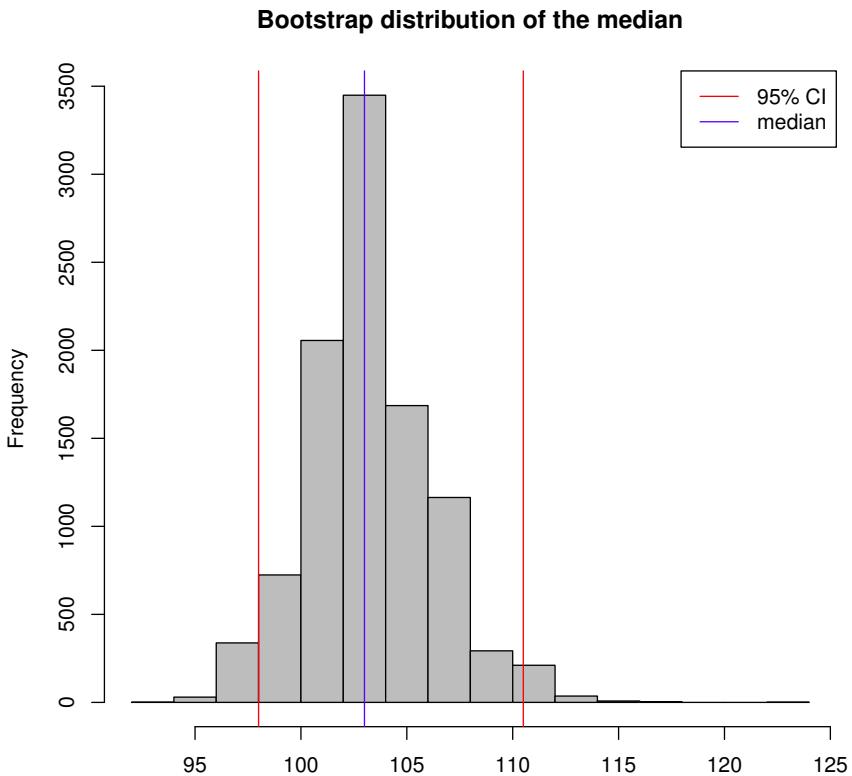
Then we can use the quantiles of these N bootstrap statistics to construct a bootstrap confidence interval.

Example: Chick weight



- central 95% of the bootstrap distribution

Example: Chick weight



- central 95% of the bootstrap distribution
- bootstrap confidence interval: $(98.0, 110.5)$

Always good?

Always good?

- your statistic T is not too weird;

Always good?

- your statistic T is not too weird;
- need a 'good' sample to start;

Always good?

- your statistic T is not too weird;
- need a 'good' sample to start;
- 'cannot' improve our point estimator.