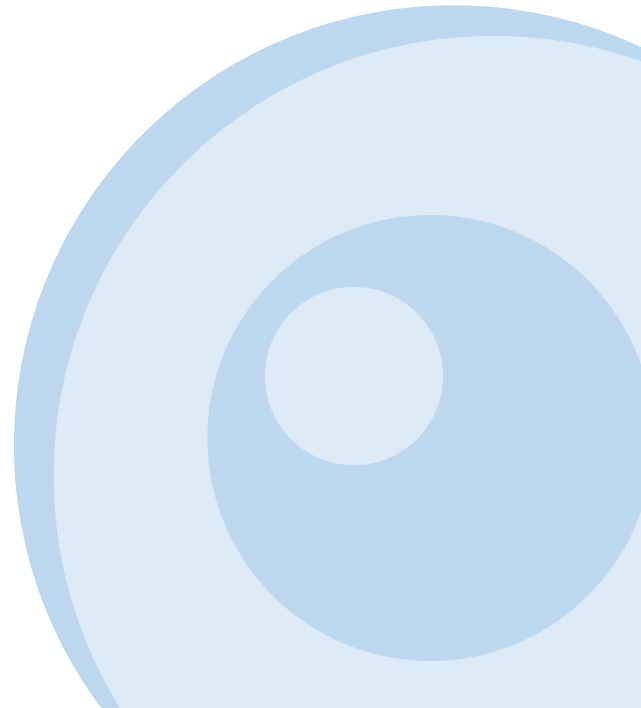


PSYC 193R:

Data Analysis & Visualization with R

Lecture 3: Parameter Estimation

Jonas S.-H. Lau [silau-et-ucsd.edu]
psyc193r.ucsd.edu



Population Parameters

Reminder: The population depends on your research question

- Central Tendency:
Mean / Median / Mode
- Dispersion:
Range / Standard Deviation
- Shape of distribution:
(Kurtosis)



Population Parameters

Reminder: The population depends on your research question

- **Central Tendency:**
Mean / Median / Mode
 - **Dispersion:**
Range / Standard Deviation
 - **Shape of distribution:**
(Kurtosis)
- What is the typical person / score in the population look like
 - How different the scores in the population are/ How precise our measurement is
 - What is the shape of the score distribution
(assumptions of many statistical tests)
- 

Parameter Estimation Problems

- E.g., we want to know how a new teaching method affects learning
 - Operational Definition of *learning*:
words remembered in 5 minutes
- That is the same as asking:
 - Central Tendency: Mean number of words students can remember
 - Dispersion: How diverse are the scores
 - Kurtosis: Does the distribution resemble a bell shape?

In psychology or related field, most of the time we're interested in *parameter estimation*

Parameter Estimation Problems

- E.g., we want to know how a new teaching method affects learning
 - Operational Definition of *learning*: # words remembered in 5 minutes
- That is the same as asking:
 - Central Tendency: Mean number of words students can remember
 - Dispersion: How diverse are the scores
 - Kurtosis: Does the distribution resemble a bell shape?

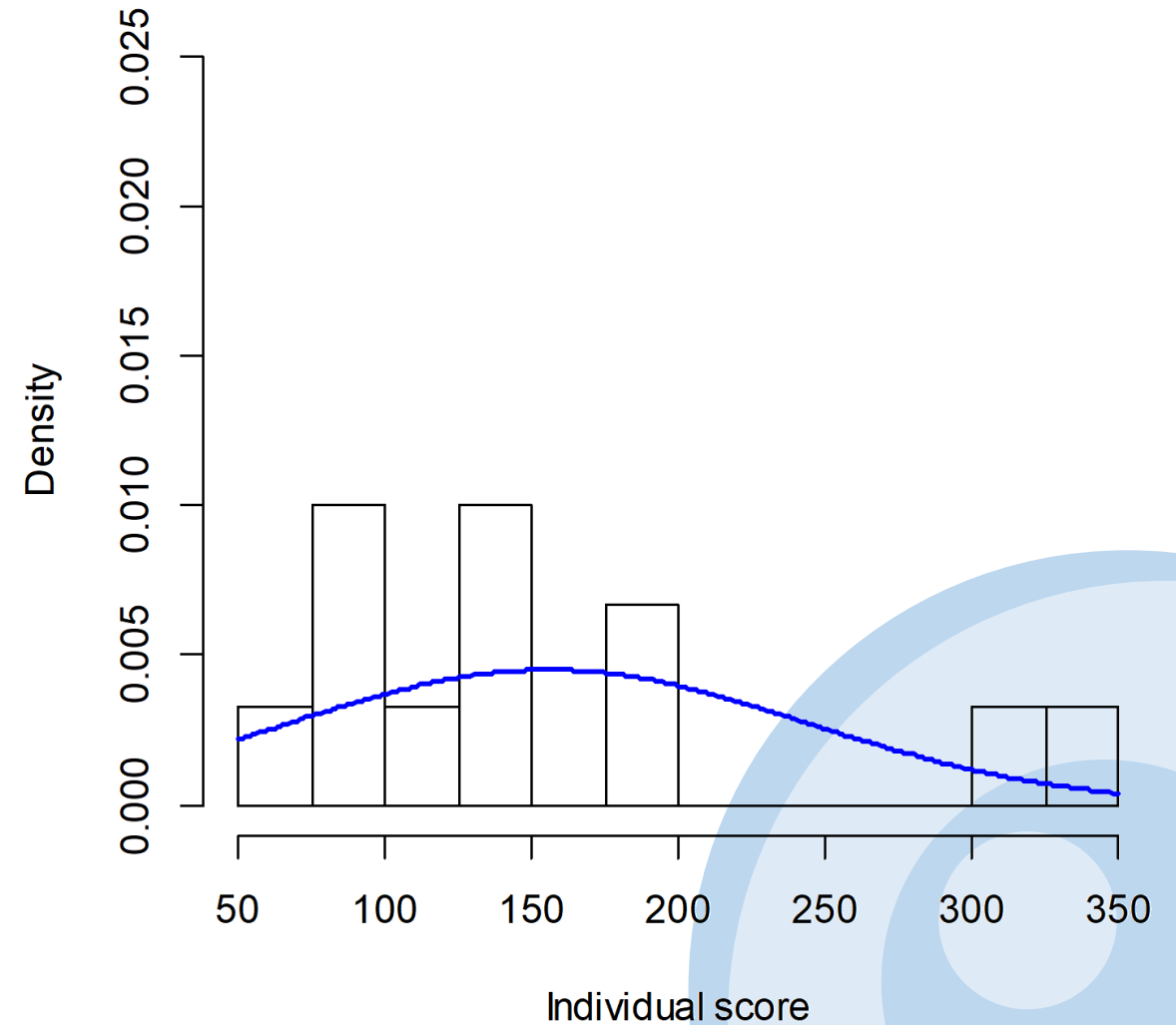
These are usually just proxies of what we want to study

Individual scores are not reliable, so we have to get a lot of scores



M&M's in a bag

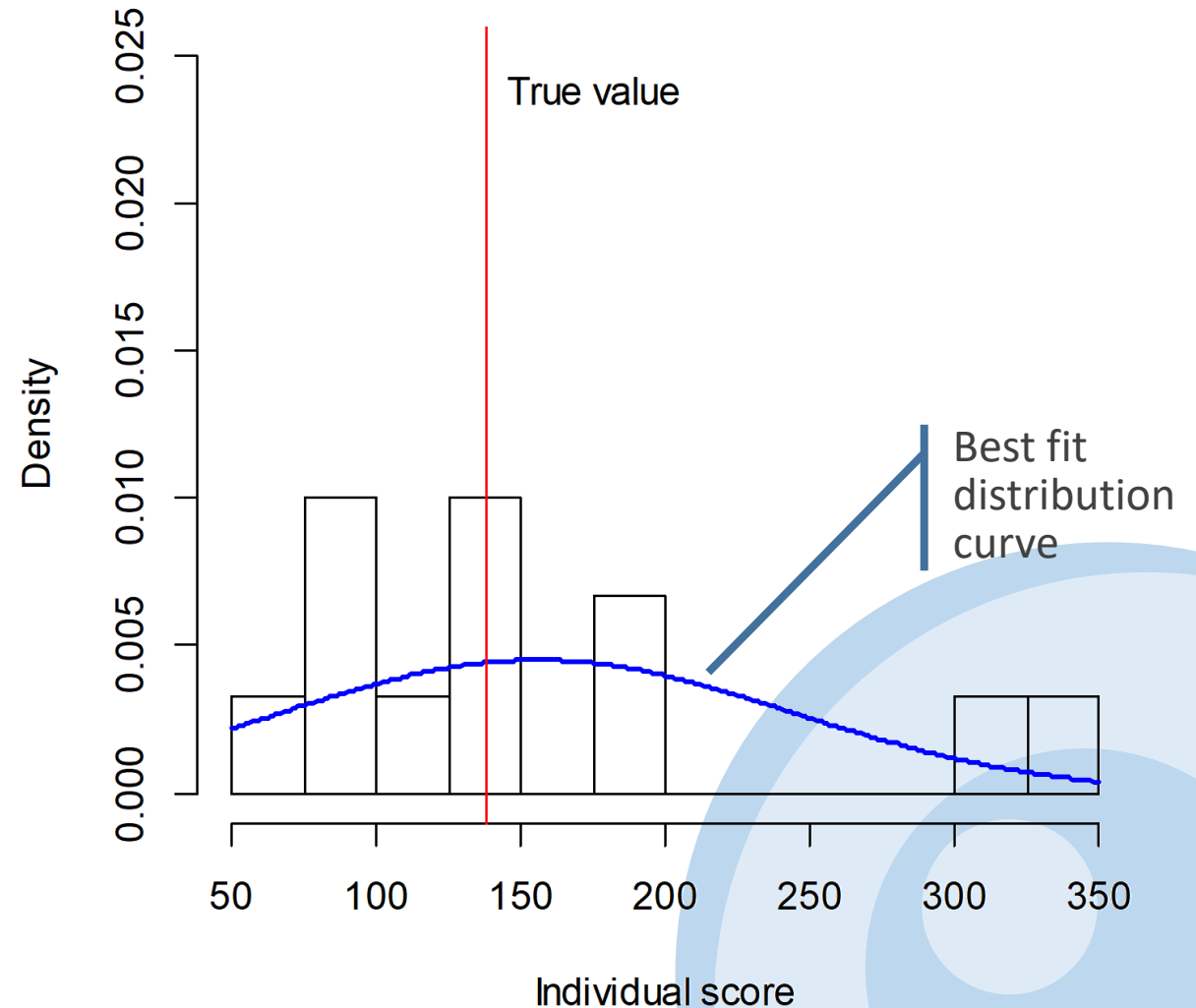
- Your Task: Estimate the number of M&M's in a bag



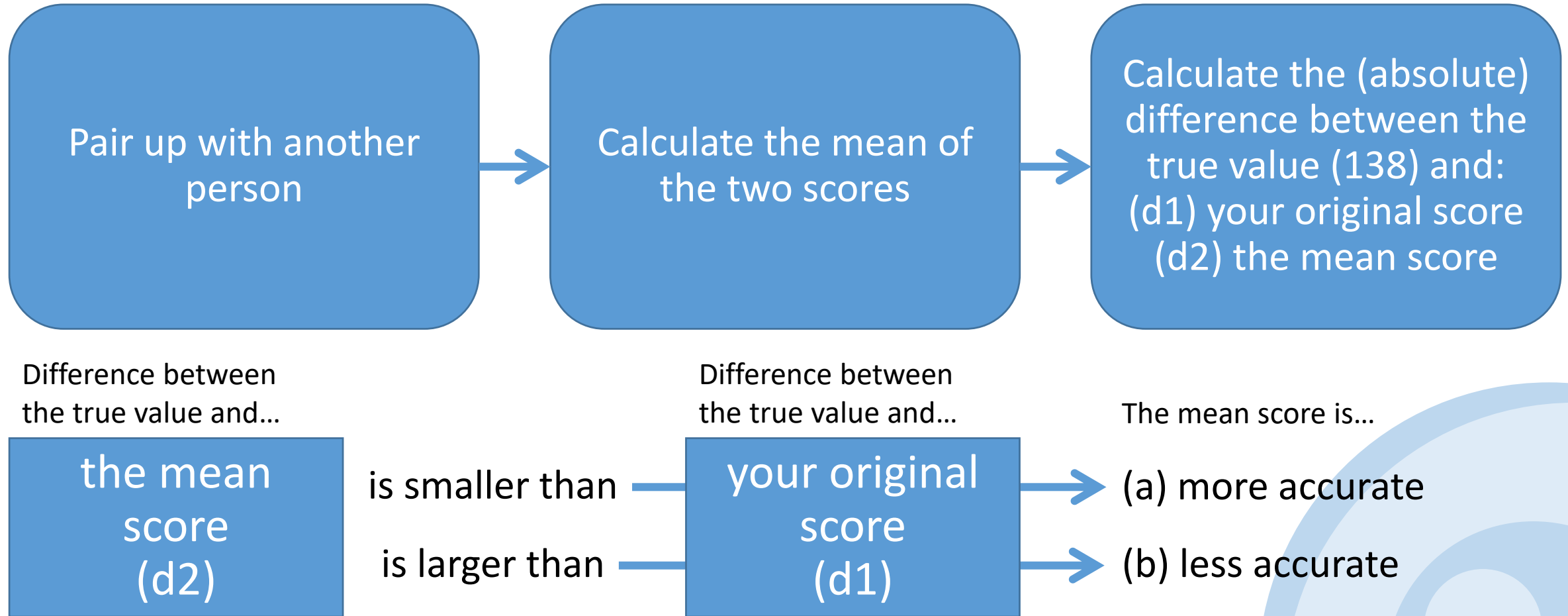
M&M's in a bag

- Your Task: Estimate the number of M&M's in a bag
- Observations:
 1. Very few people are accurate (dispersion is large, $\sigma = 88.27$)
 2. Mean score ($\mu=155.5$) is close to the true value (138)

Law of Large Numbers:
The more "estimates" you average, the closer you will get to the population parameter

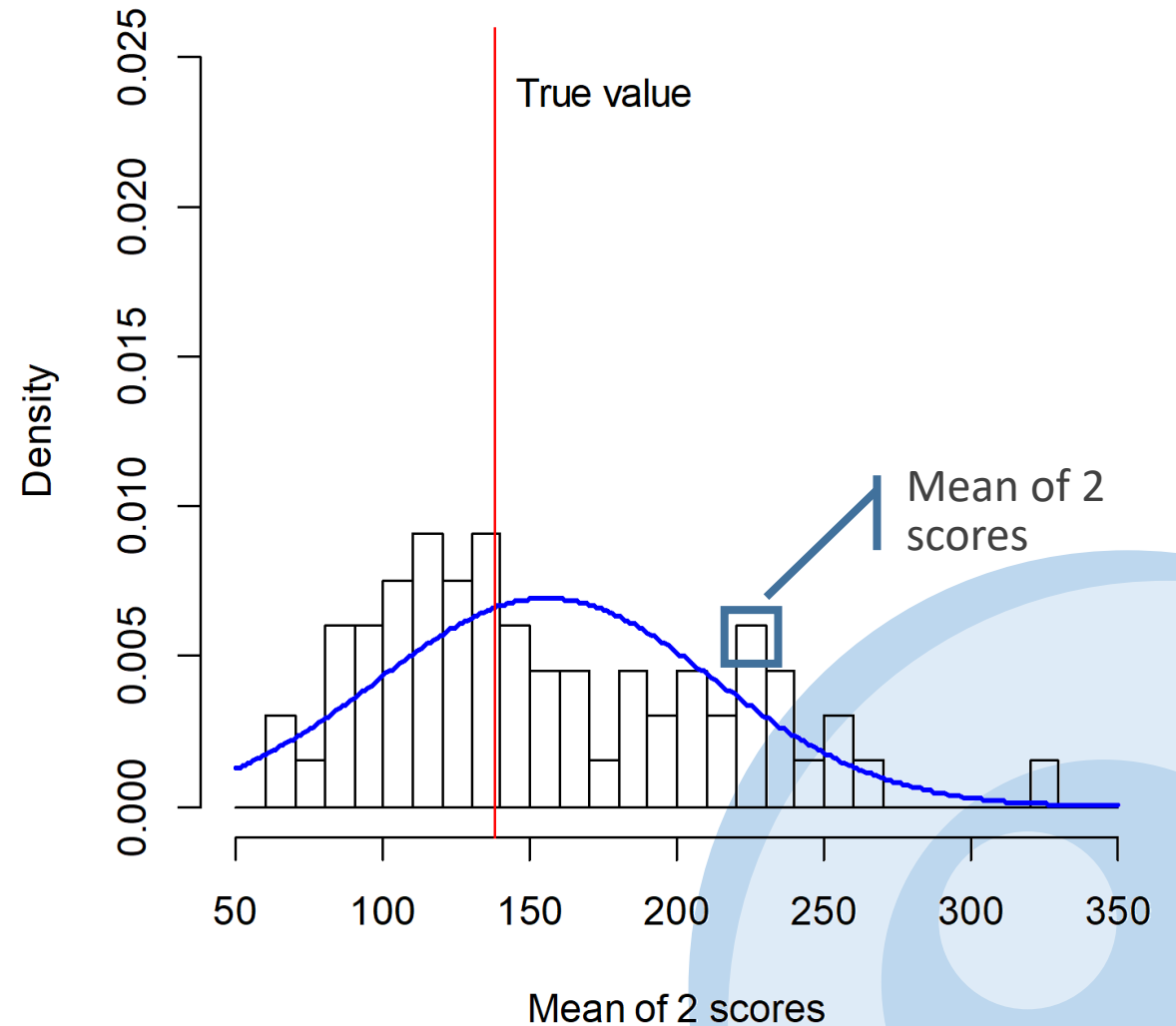


Class exercise: Averaging 2 scores



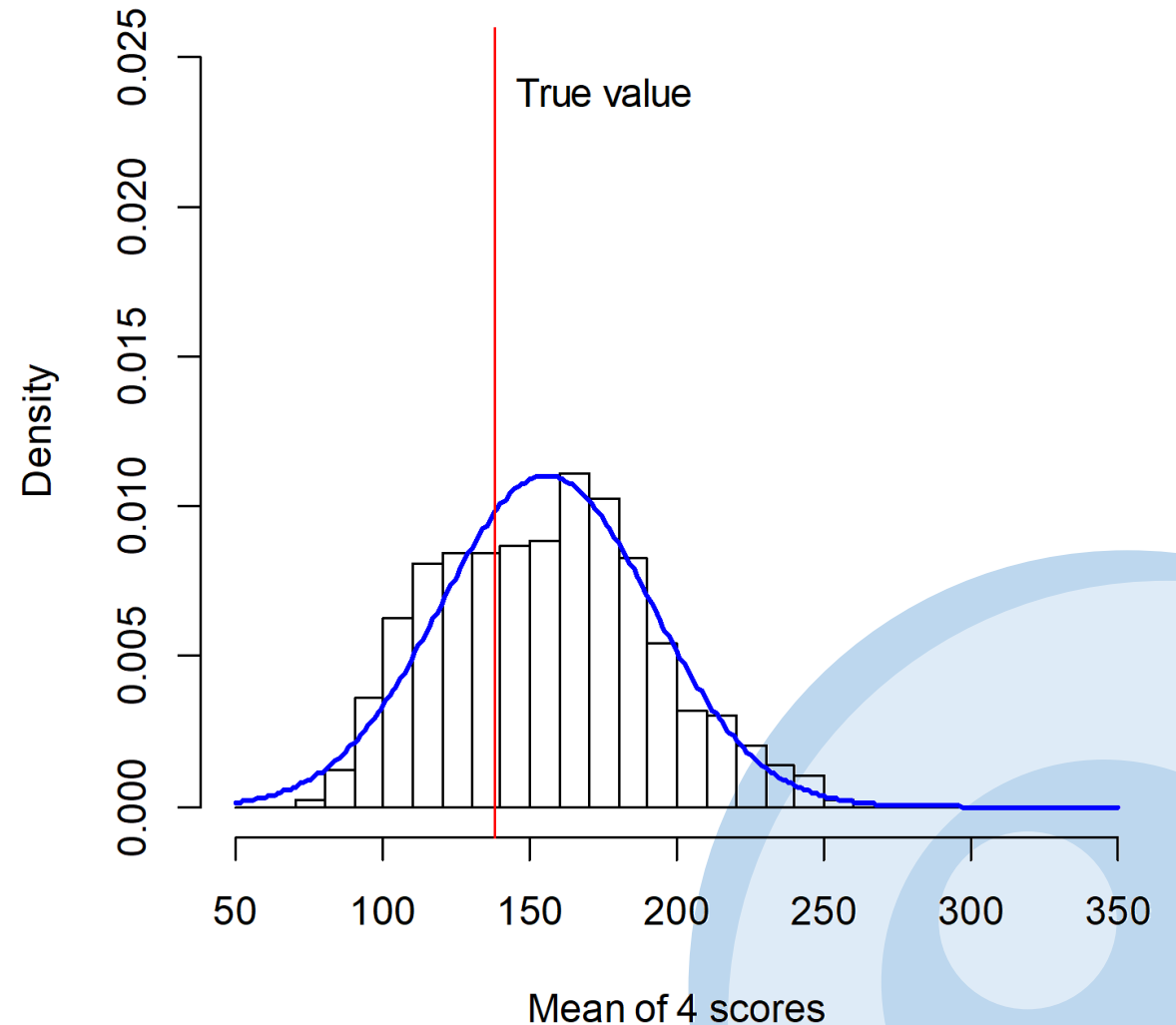
M&M's in a bag (Sampling Distribution of 2 scores)

- Your Task: Estimate the number of M&M's in a bag
- Forming a Sampling distribution:
 - Picking 2 scores from the original distribution
 - Calculate the mean
 - Repeat and plot the means
- Observations:
 1. Mean of the distribution does not change
 2. Dispersion is smaller



M&M's in a bag (Sampling Distribution of 4 scores)

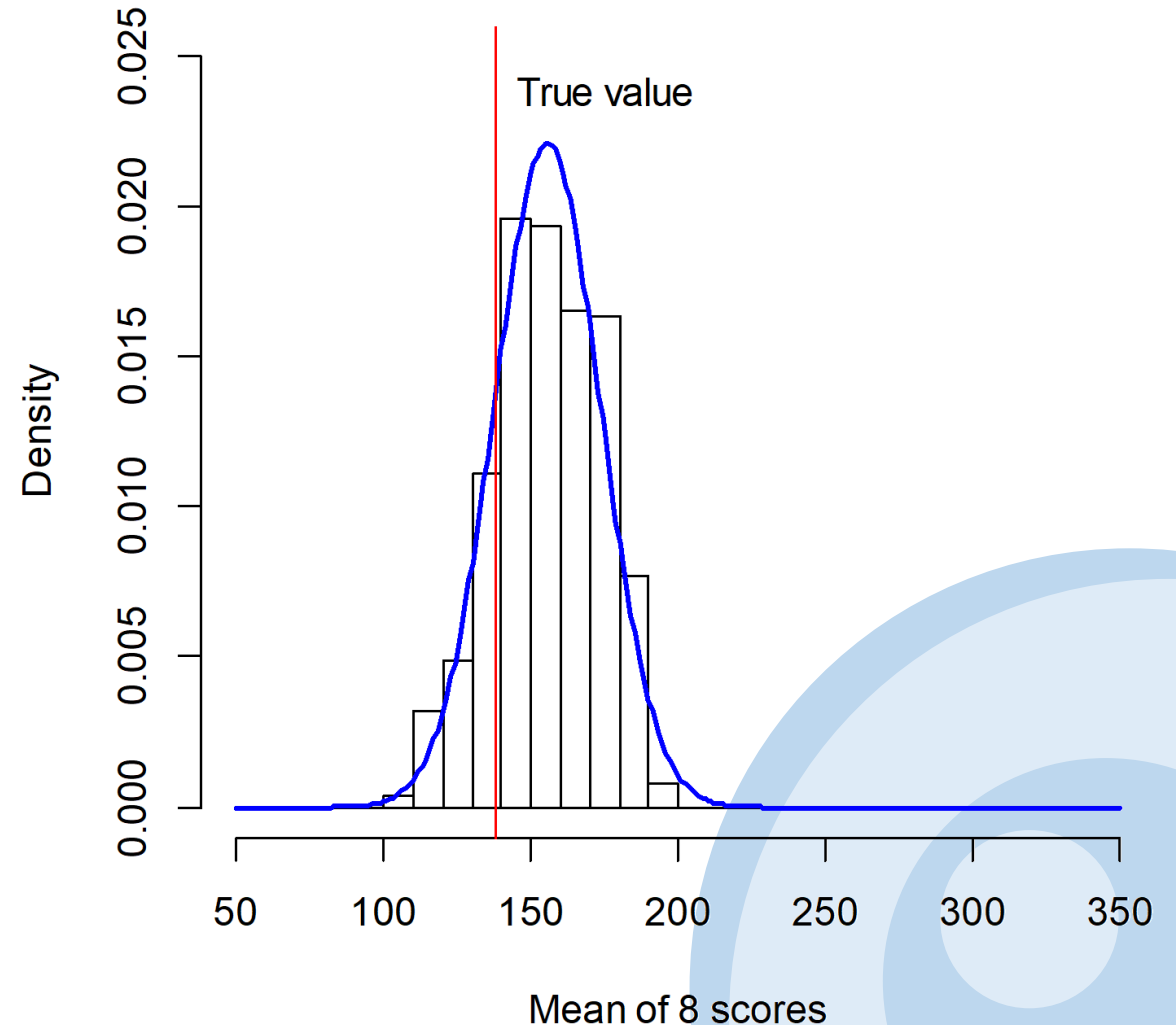
- Your Task: Estimate the number of M&M's in a bag
- Sampling distribution:
 1. Mean of the distribution does not change
 2. Dispersion is smaller
 3. Distribution is approaching "Normal"



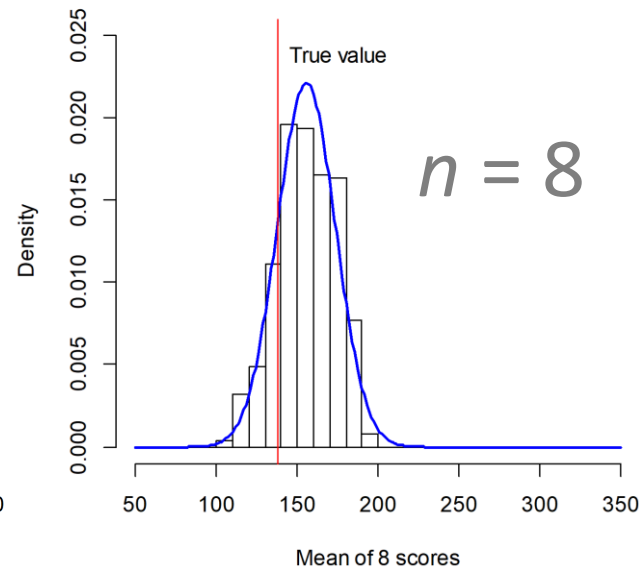
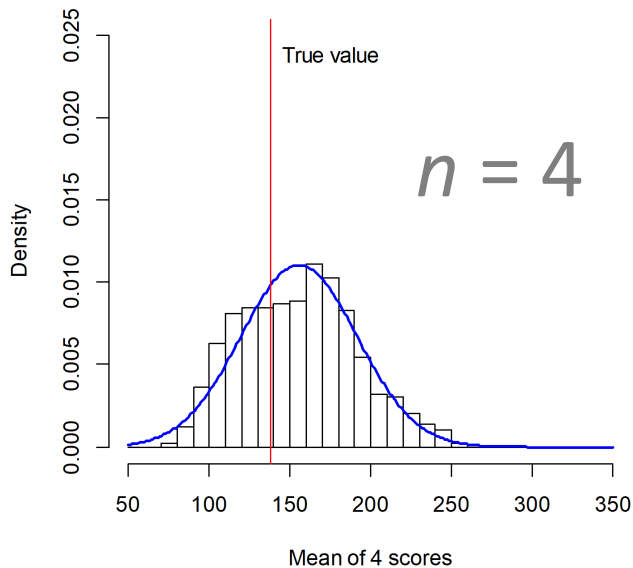
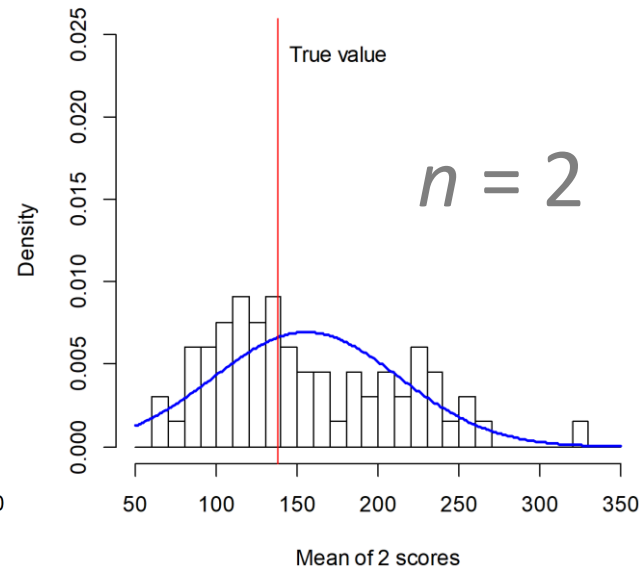
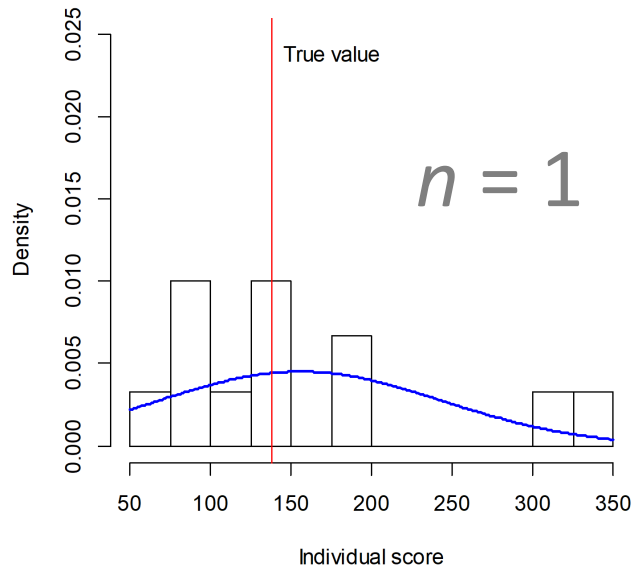
M&M's in a bag (Sampling Distribution of 8 scores)

- Your Task: Estimate the number of M&M's in a bag
- Sampling distribution:
 1. Mean of the distribution does not change
 2. Dispersion is smaller (you're more confident with your estimate)
 3. Distribution is approaching "Normal" (good for many statistical tests)

Central Limit Theorem



Central Limit Theorem



- Mean of the distribution does not change

$$\mu_M = \mu$$

- Dispersion is smaller (standard error $\downarrow$)

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

- Distribution is approaching "Normal"
(if $n > 30$, the distribution is non-distinguishable from normal distribution)

Implications

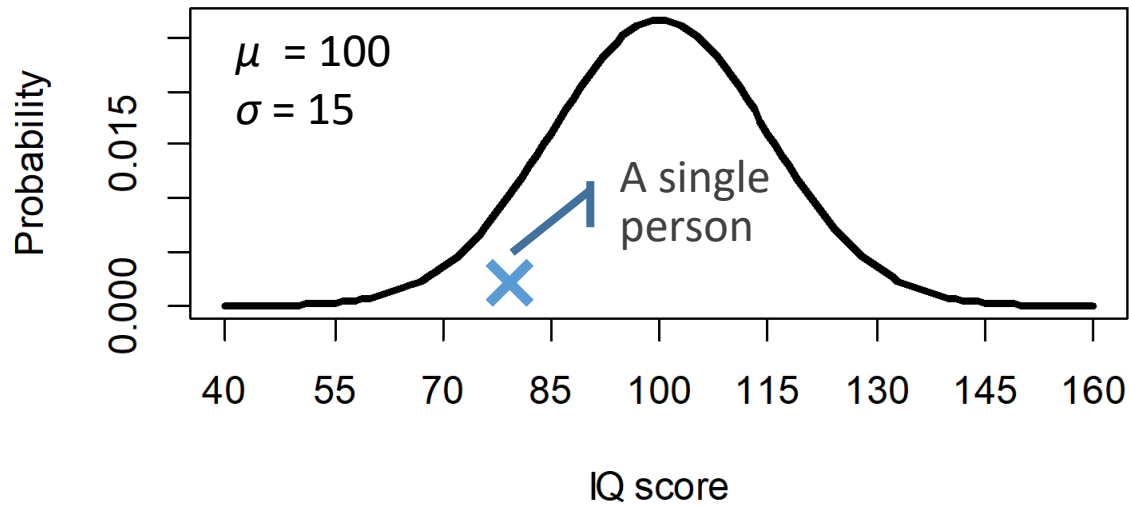
Law of Large Numbers:

- The more "estimates" you average, the closer you will get to the population parameter
- When you are estimating the population parameters, get multiple measures
 - You test multiple people to estimate the population parameter
 - You have multiple questions in a test for each participant

Central Limit Theorem:

- Dispersion is smaller
- Distribution is approaching "Normal"
- When you are comparing a sample mean vs a population, convert population distribution into a *sampling* distribution prior to testing
- Use a sample size >30 to get a "normal" sampling distribution (if the original population distribution doesn't follow a "normal" shape)

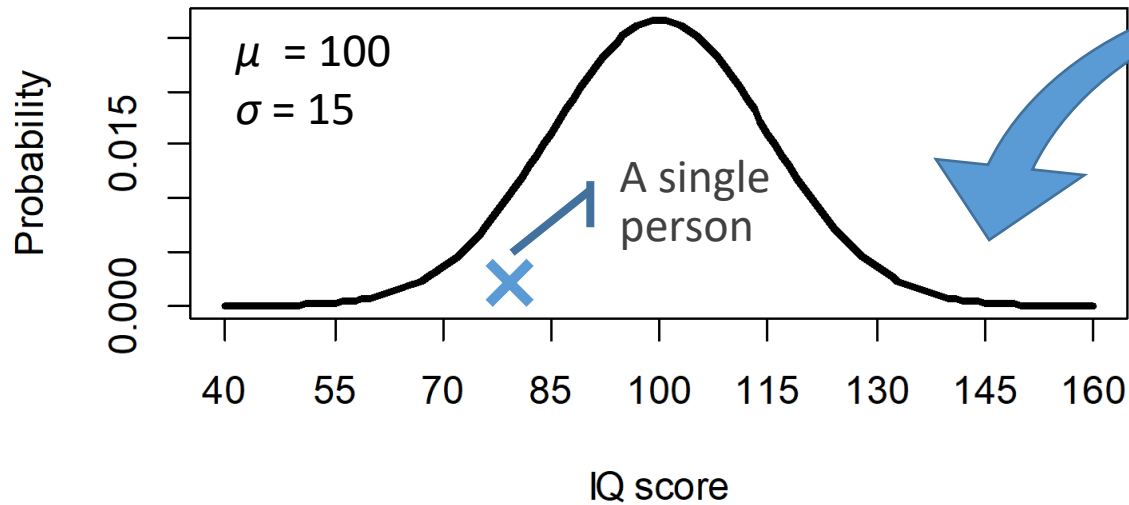
Implications: Testing an IQ drug



Law of Large Numbers:

1. When you are estimating the population parameters (people's IQ *with* the drug), get multiple measures

Implications: Testing an IQ drug

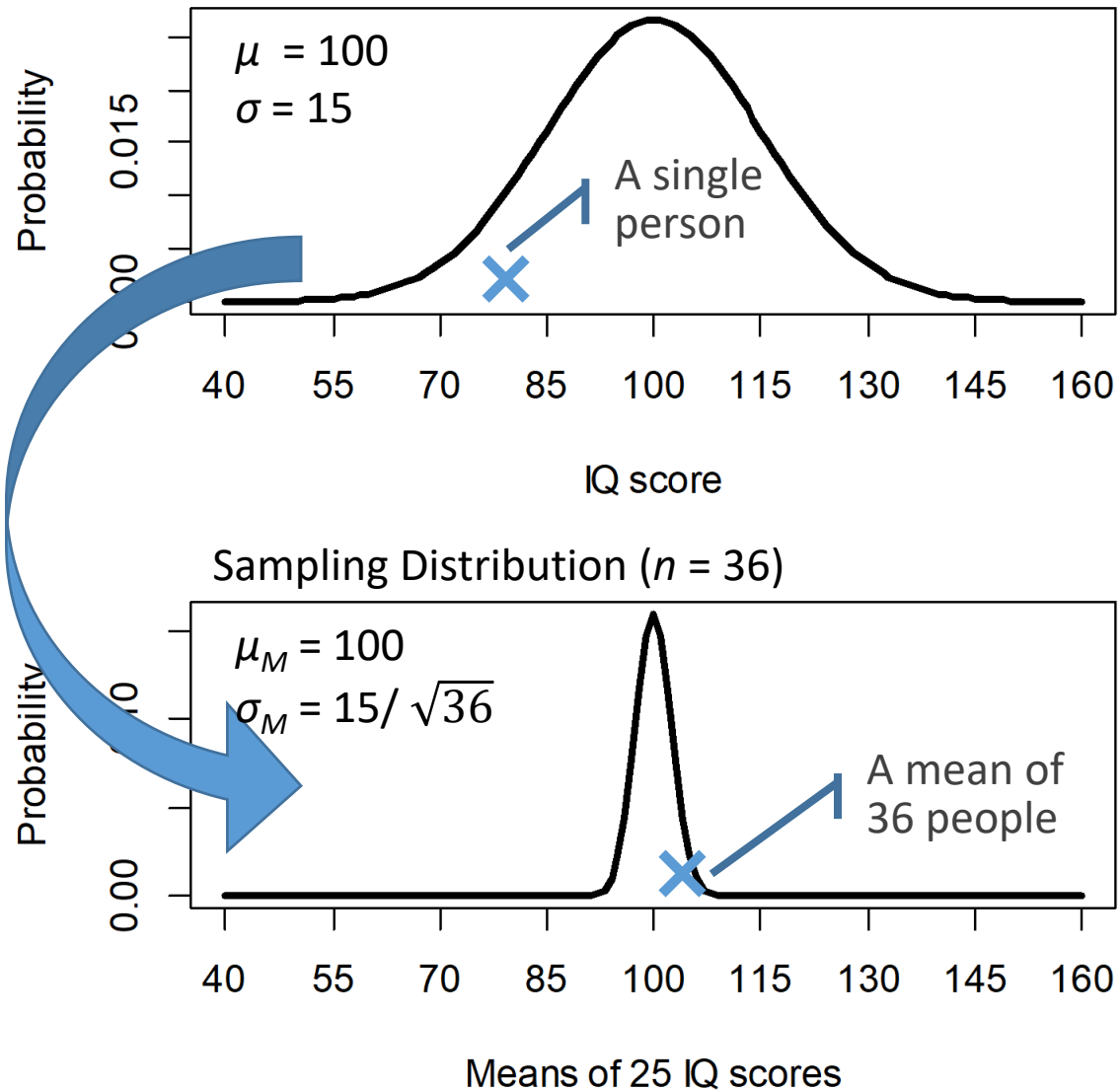


Law of Large Numbers:

1. When you are estimating the population parameters (people's IQ *with* the drug), get multiple measures

Say, $n = 36$

Implications: Testing an IQ drug



Law of Large Numbers:

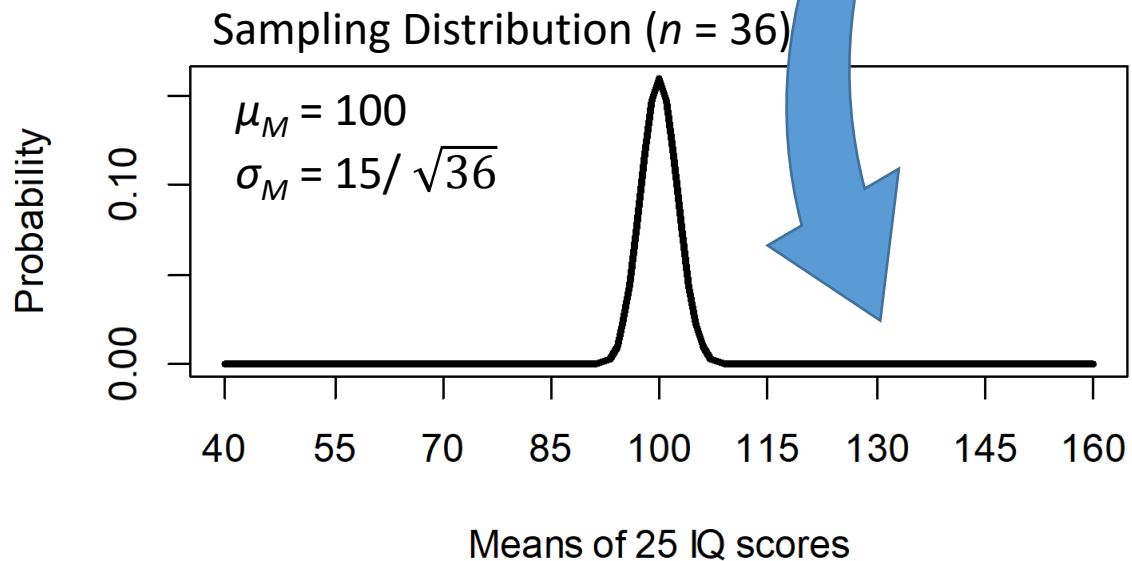
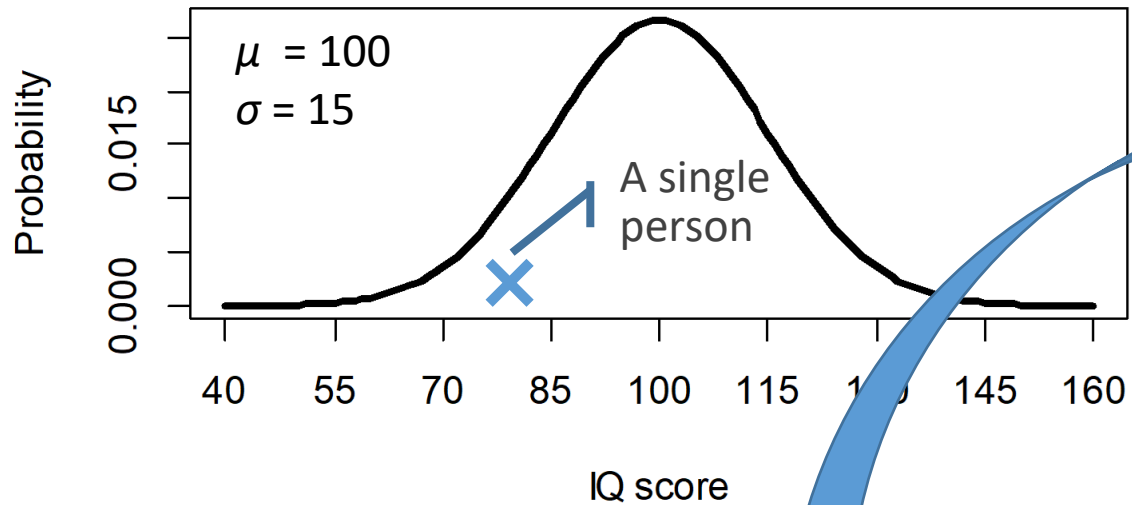
1. When you are estimating the population parameters (people's IQ *with* the drug), get multiple measures

Say, $n = 36$ ●

Central Limit Theorem:

2. Convert population distribution into a sampling distribution
3. Use a sample size >30 to get a "normal" sampling distribution

Implications: Testing an IQ drug



Law of Large Numbers:

1. When you are estimating the population parameters (people's IQ *with* the drug), get multiple measures

Say, $n = 36$

Central Limit Theorem:

2. Convert population distribution into a sampling distribution
3. Use a sample size >30 to get a "normal" sampling distribution