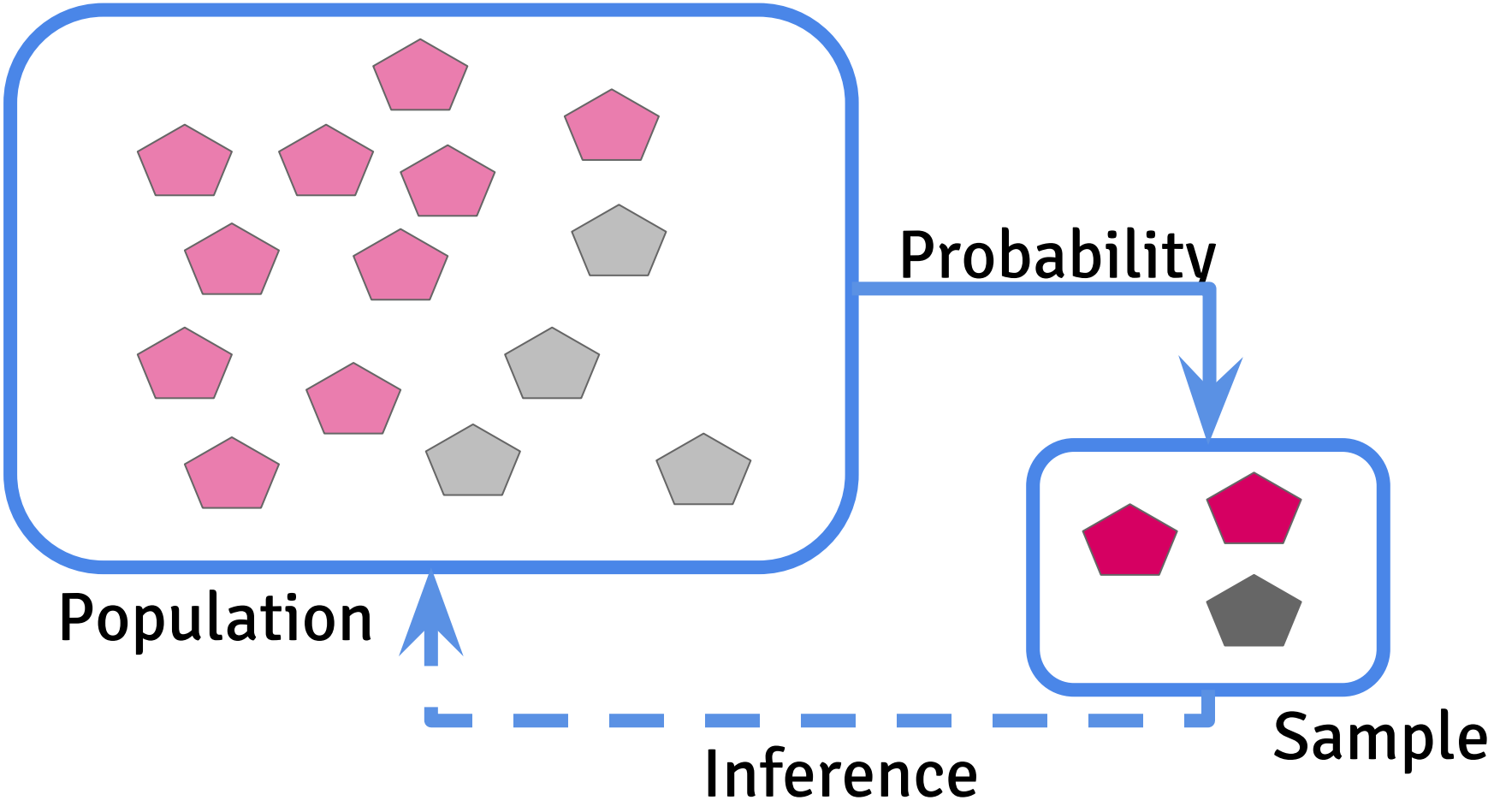
Inference

Jeff Leek

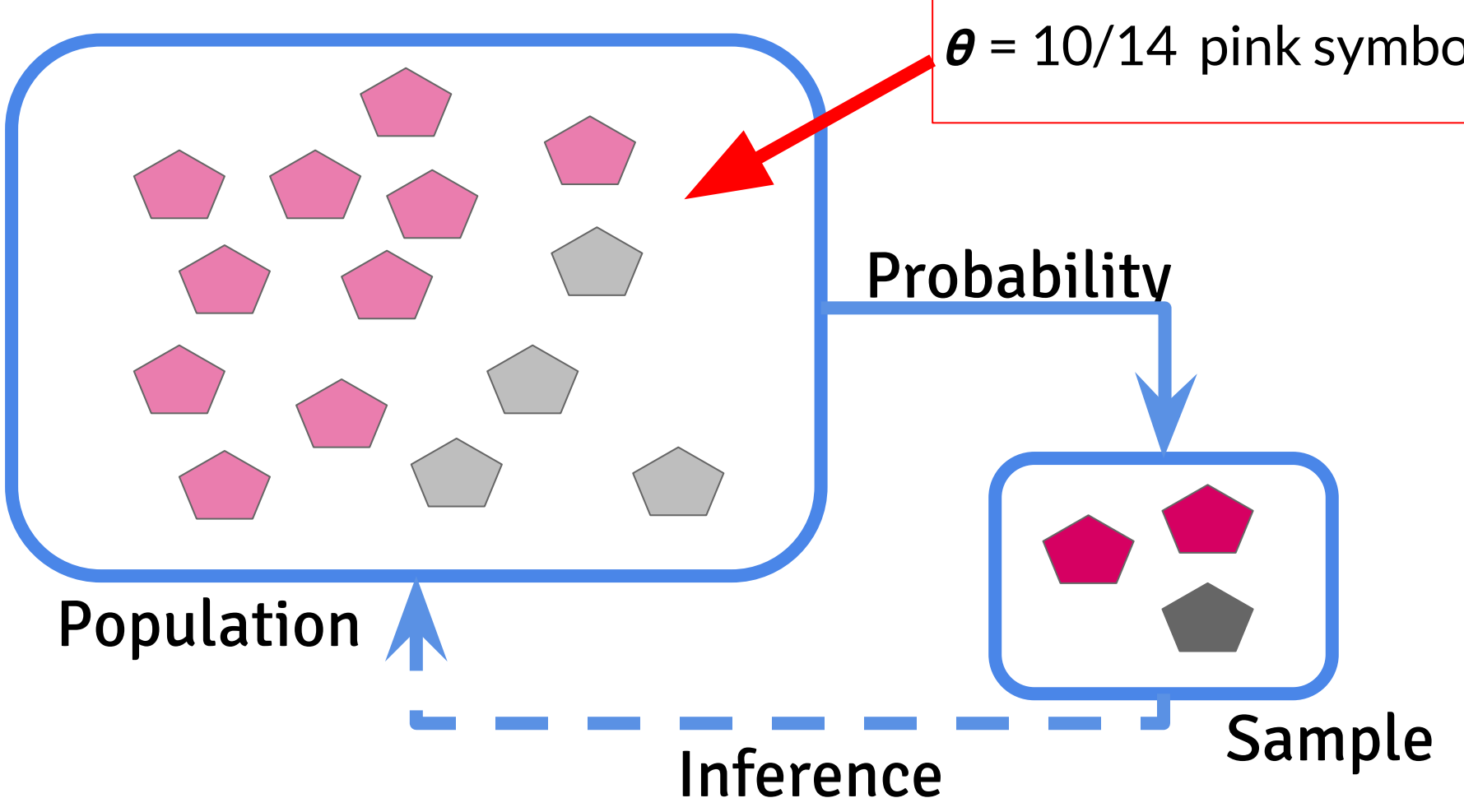
@jtleek

www.jtleek.com

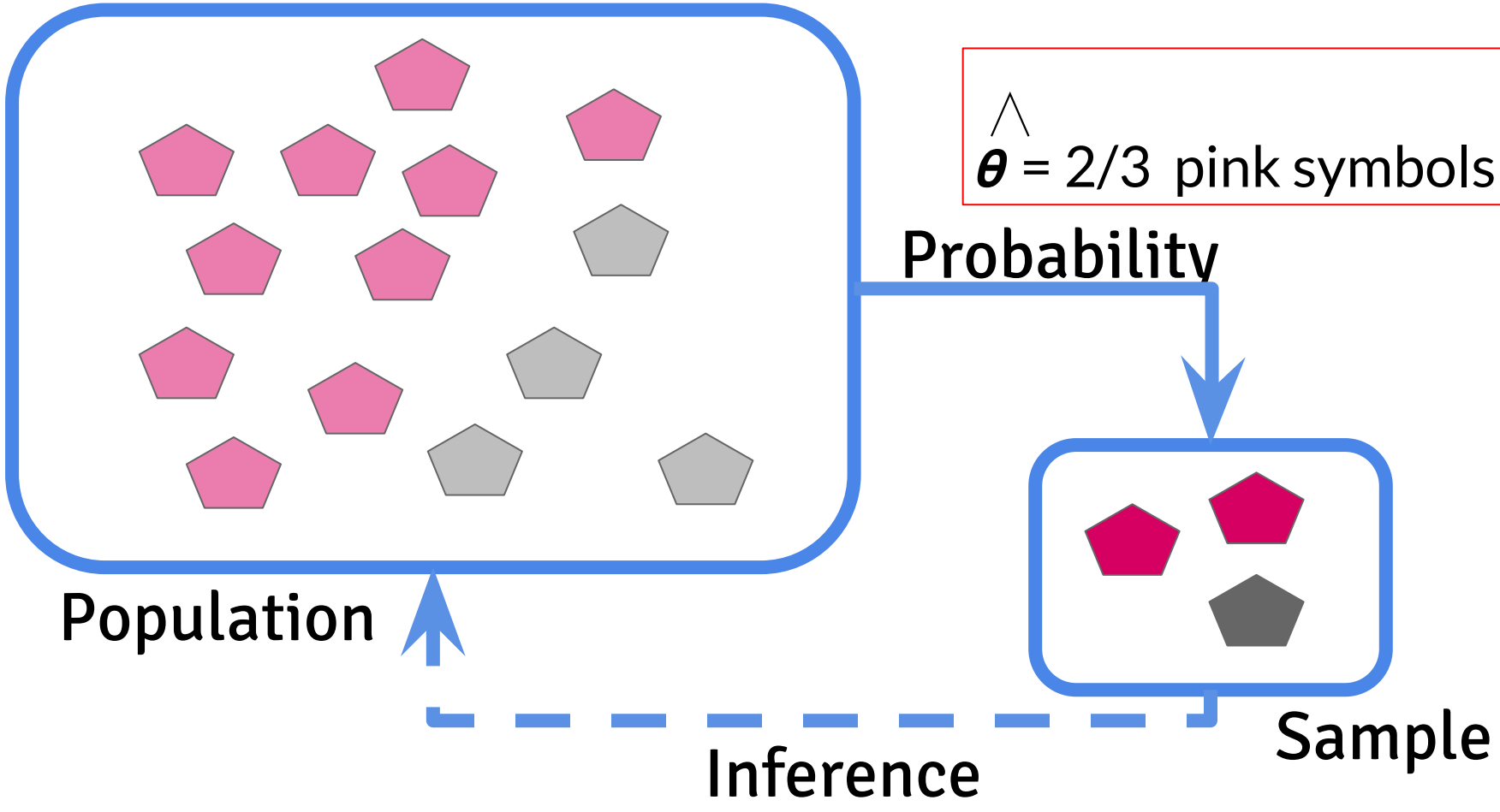
Central dogma of statistics


Probability
Population
Sample
Inference

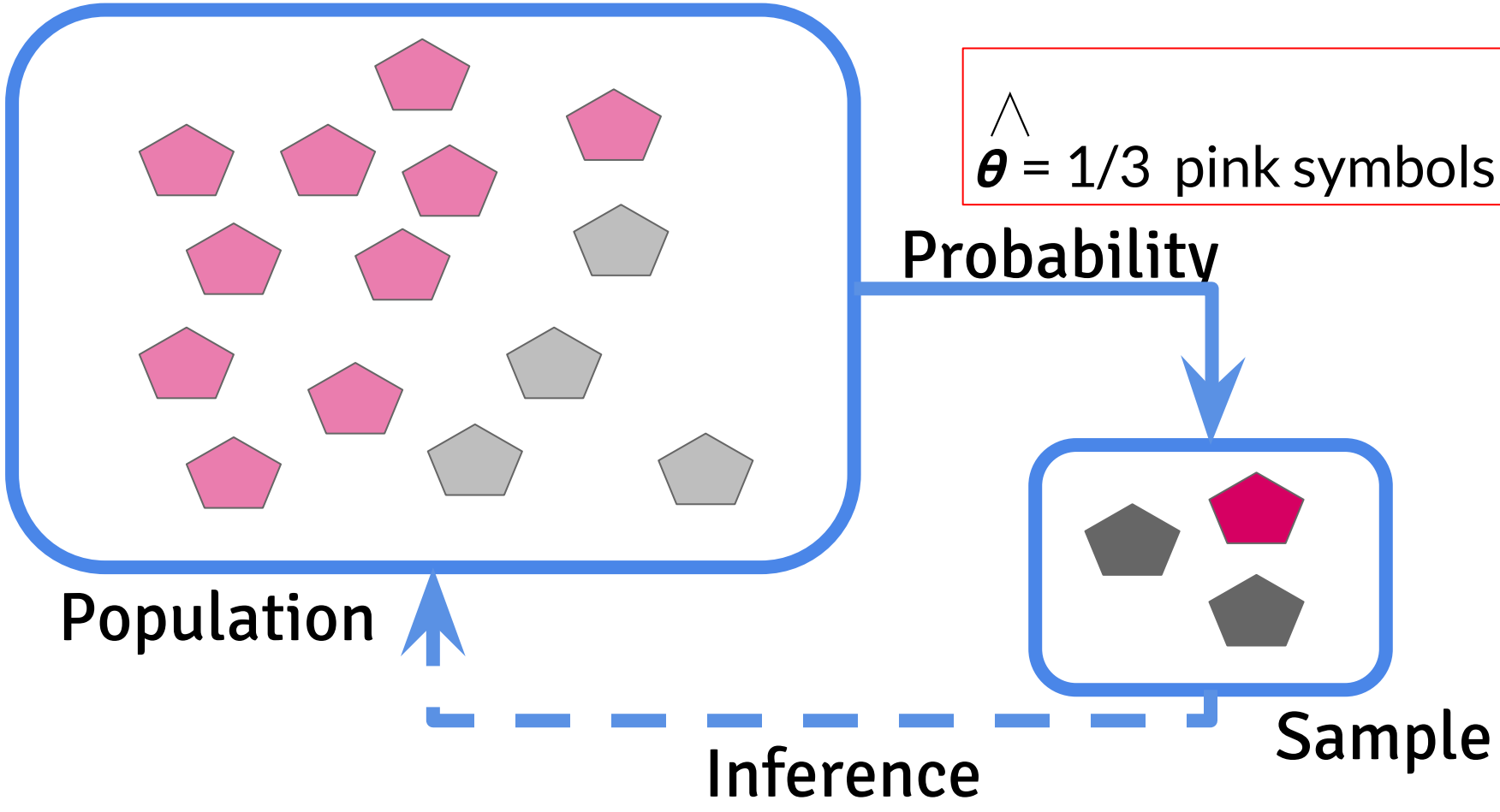
Parameters are characteristics of the population


θ = 10/14 pink symbols
Probability
Population
Sample
Inference

We estimate population parameters with the data

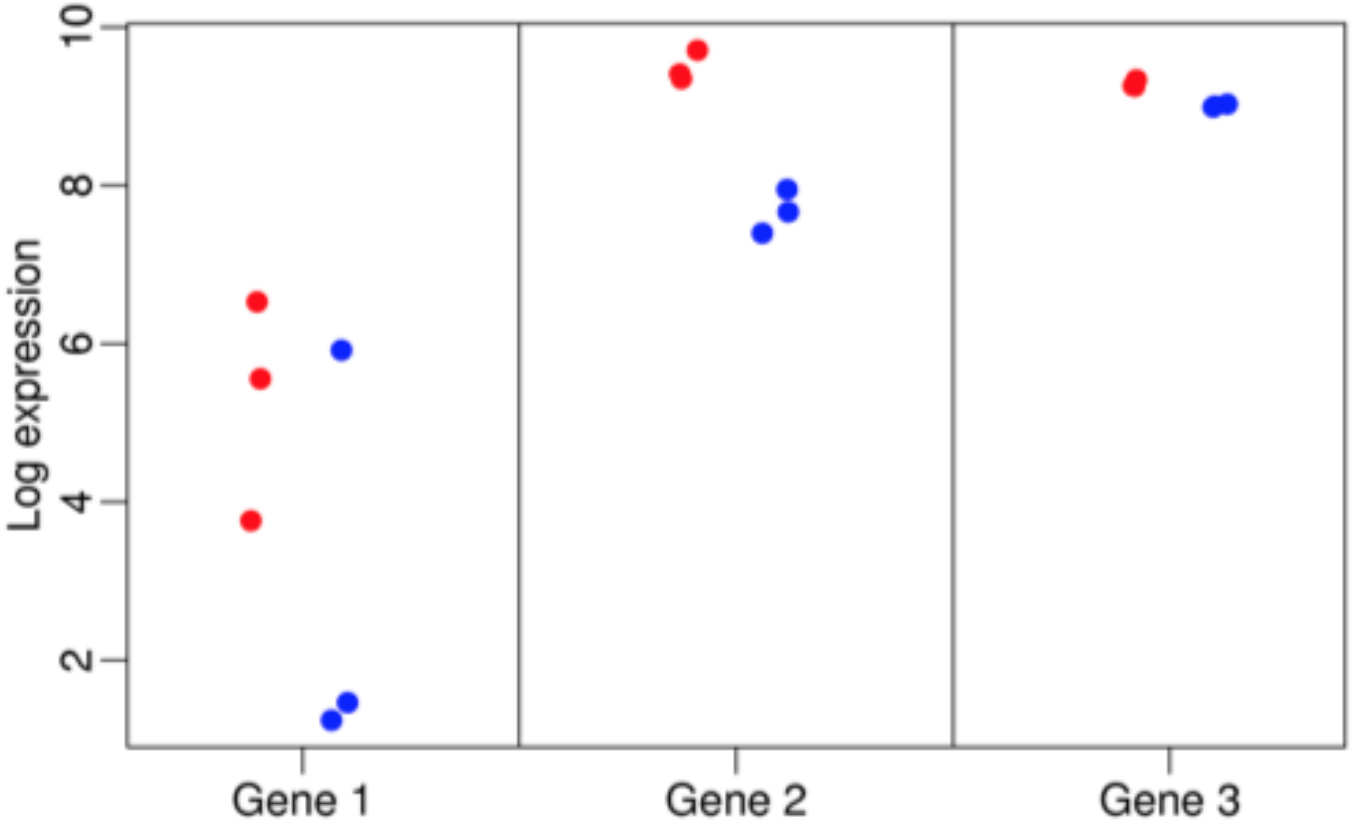

$\hat{\theta}$ = 2/3 pink symbols
Probability
Population
Sample
Inference

But in another sample we might get something different


$\hat{\theta}$ = 1/3 pink symbols
Probability
Population
Sample
Inference

Goal : what is population quantity and how certain are we?

Slide courtesy: Ingo Ruczinski

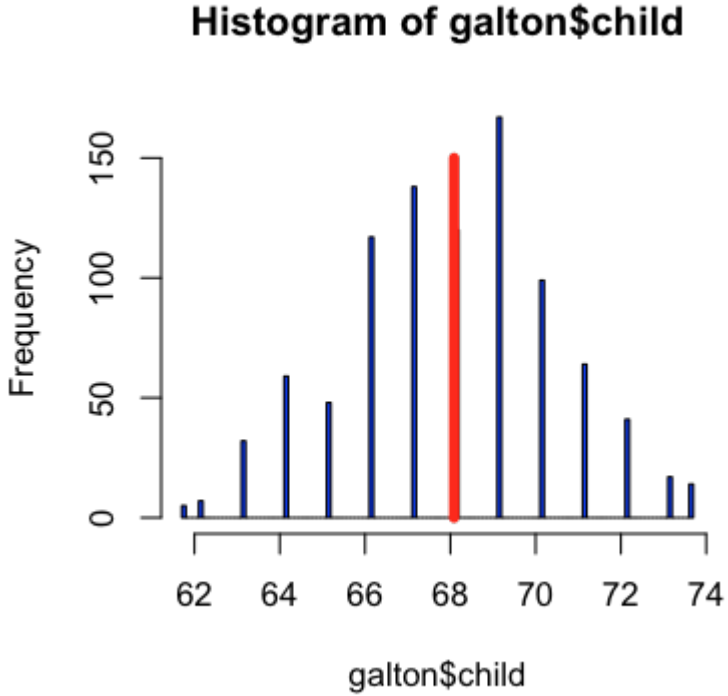


$$\bar{X} = \frac{1}{M}\sum_{i=1}^{M} X_i$$

$$s_X^2 = \frac{1}{M-1}\sum_{i=1}^{M} (X_i - \bar{X})^2$$

$$CI = (\bar{X} - c\frac{s_x}{\sqrt{n}}, \bar{X} + c\frac{s_x}{\sqrt{n}})$$

$$\Pr(\theta \in CI) = f(c)$$



Notes and further reading

- Inference is a whole class (no joke): https://www.coursera.org/course/statinference
- Basic thing to keep in mind is how do you
  - estimate
  - quantify uncertainty