T-distribution

Overview of Today

Overview of Today

- Guinness

Overview of Today

- Guinness

- But first…a quick review of the Z-distribution

Overview of Today

- Guinness

- But first…a quick review of the Z-distribution

- T-distribution: comparing sample means, small sample sizes

Review of Z-distribution

A Quick Review of the Z- (Normal) Distribution and Z-Tests

- The Z-distribution can be used to test differences in means

A Quick Review of the Z-Distribution and Z-Tests

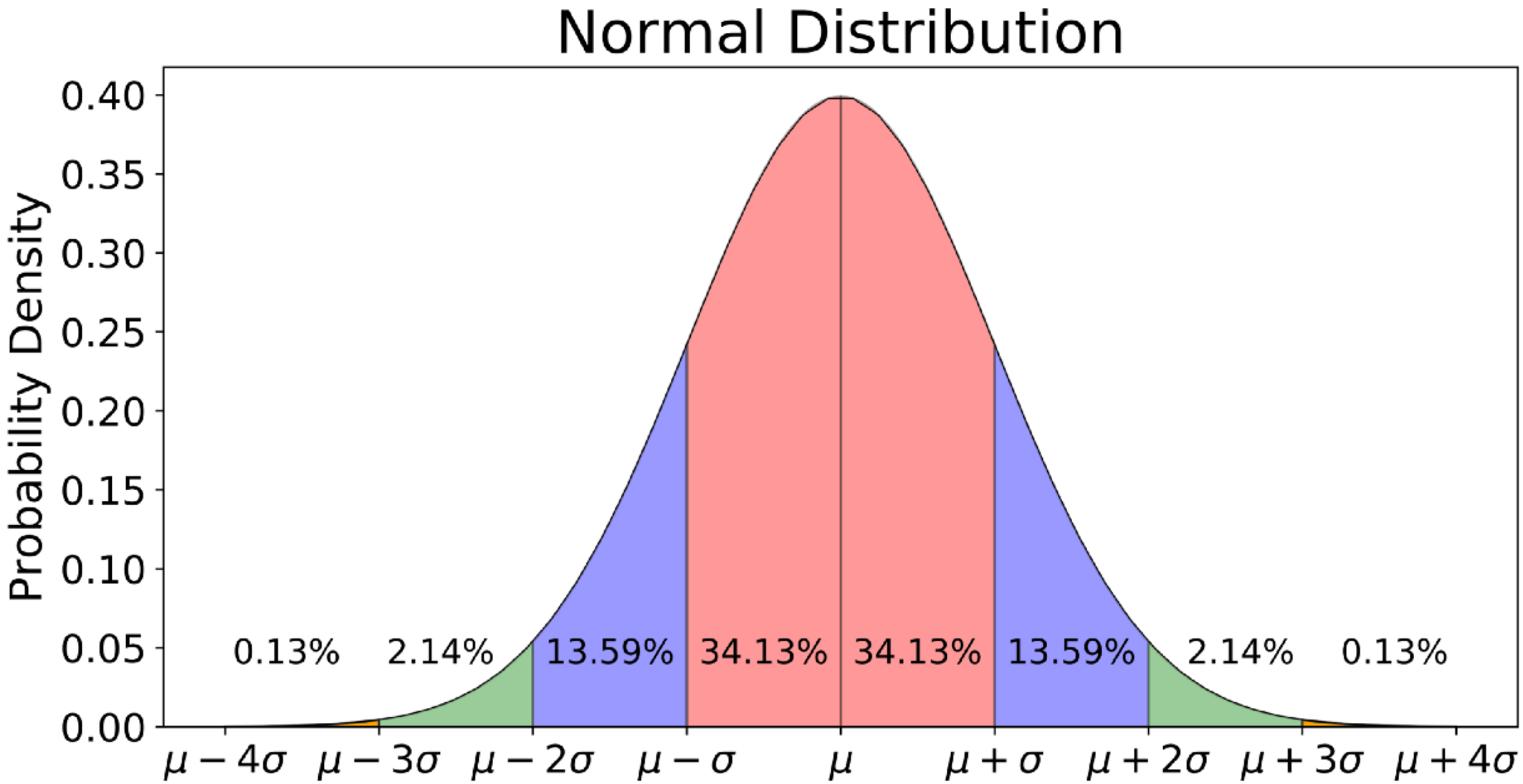
- The Z-distribution can be used to test differences in means
  - *When the population parameters are known*

A Quick Review of the Z-Distribution and Z-Tests

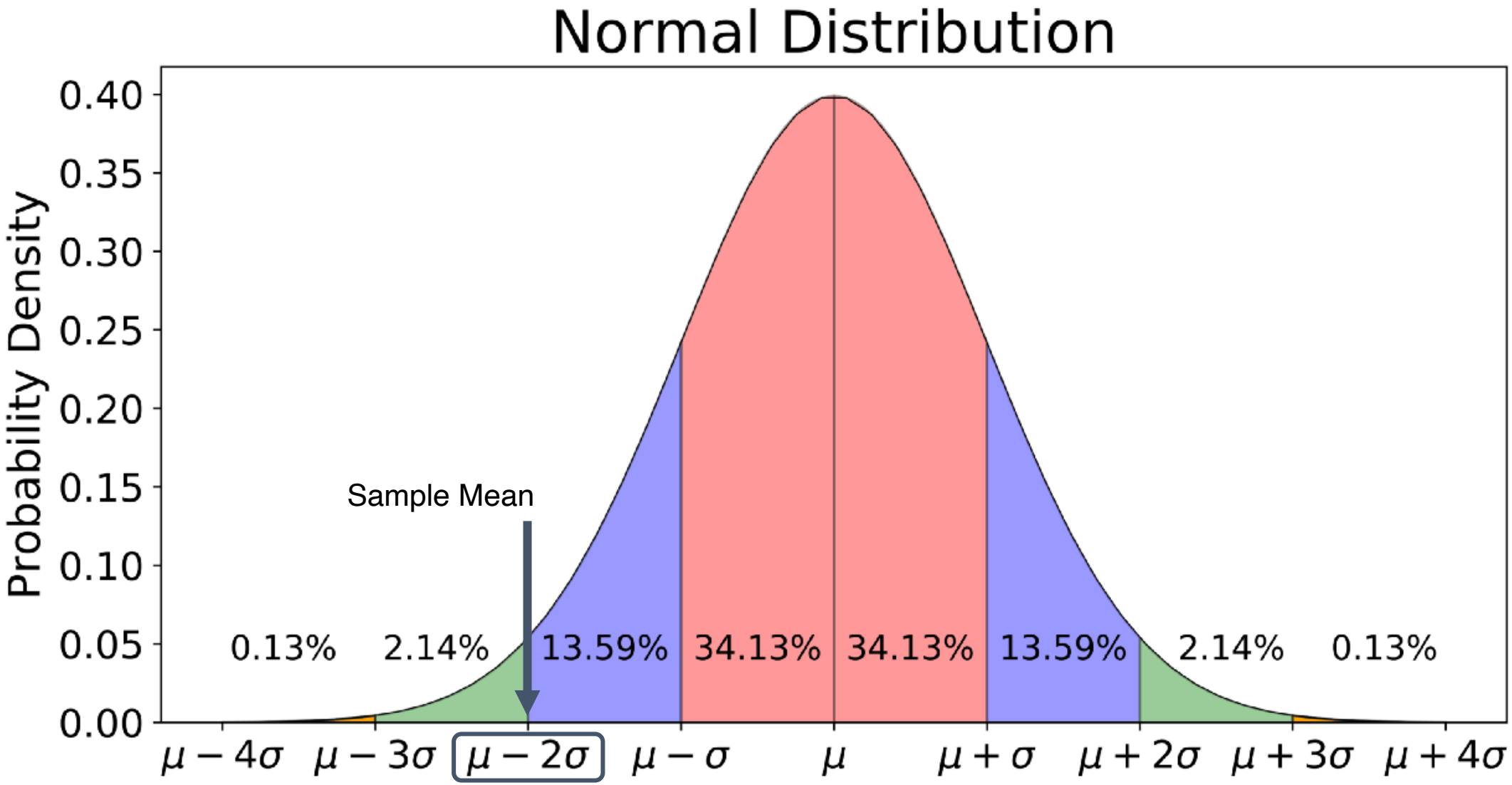
- The Z-distribution can be used to test differences in means
  - *When the population parameters are known*

- Known probability of a sample mean ($\bar{x}$) given the population mean (μ) and **standard deviation (σ) / variance (σ^2)**

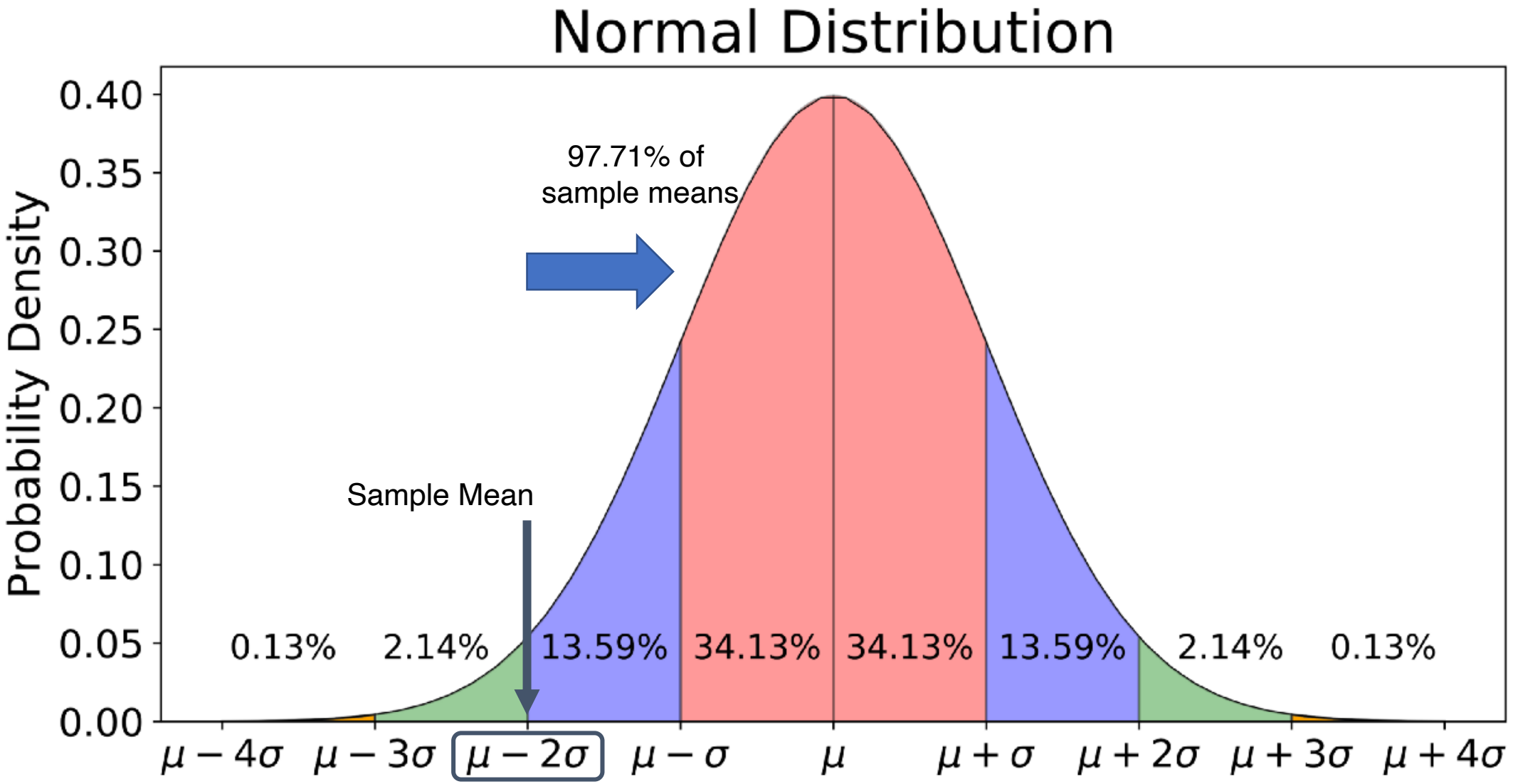
A Quick Review of the Z-Distribution and Z-Tests



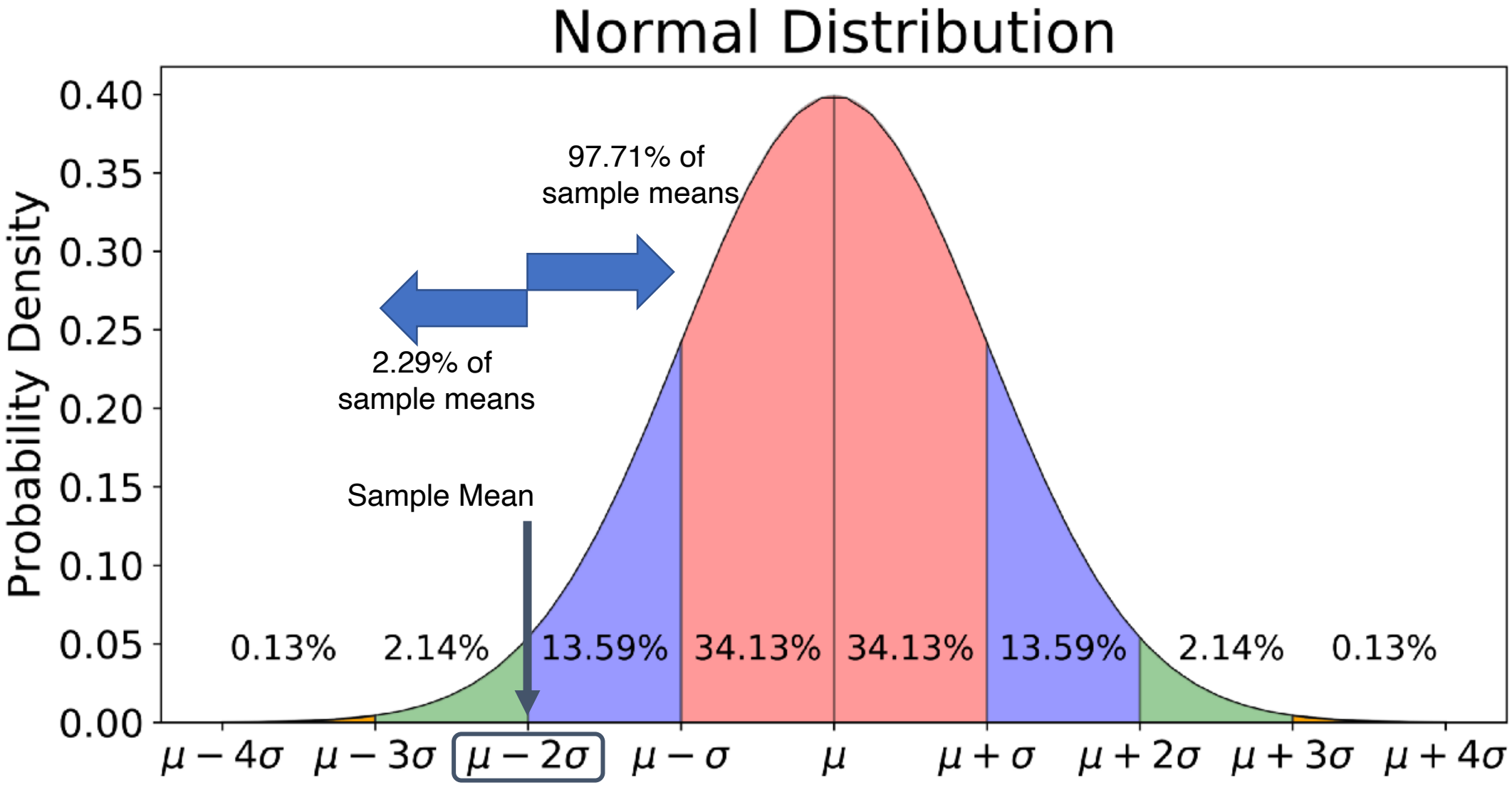
A Quick Review of the Z- (Normal) Distribution and Z-Tests



A Quick Review of the Z- (Normal) Distribution and Z-Tests



A Quick Review of the Z- (Normal) Distribution and Z-Tests



A Quick Review of the Z-Distribution and Z-Tests

- The Z-test relies on knowing the population parameters that describe the Z-distribution
  - How often do we really know population parameters?

A Quick Review of the Z-Distribution and Z-Tests

- The Z-test relies on knowing the population parameters that describe the Z-distribution
  - How often do we really know population parameters?

- We need to account for the additional uncertainty in our estimates when doing hypothesis tests

A Quick Review of the Z-Distribution and Z-Tests

- The Z-test relies on knowing the population parameters that describe the Z-distribution
  - How often do we really know population parameters?

- We need to account for the additional uncertainty in our estimates when doing hypothesis tests

Variance Formulas:

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N-1}$$

A Quick Review of the Z-Distribution and Z-Tests

- The Z-test relies on knowing the population parameters that describe the Z-distribution
  - How often do we really know population parameters?

- We need to account for the additional uncertainty in our estimates when doing hypothesis tests

Variance Formulas:

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N-1}$$

Student’s T-distribution

T-distribution: History and Motivation

- Original developed by William Gosset (“Student”)

T-distribution: History and Motivation

- Original developed by William Gosset (“Student”)

- Used when we **do not** know the population mean & standard deviation
  - How much variability do we expect in the means?

T-distribution: History and Motivation

- Original developed by William Gosset (“Student”)

- Used when we **do not** know the population mean & standard deviation
  - How much variability do we expect in the means?

- T-tests to compare sample means
  - Particularly useful in small sample sizes

T-distribution: History and Motivation

- Original developed by William Gosset (“Student”)

- Used when we **do not** know the population mean & standard deviation
  - How much variability do we expect in the means?

- T-tests to compare sample means
  - Particularly useful in small sample sizes

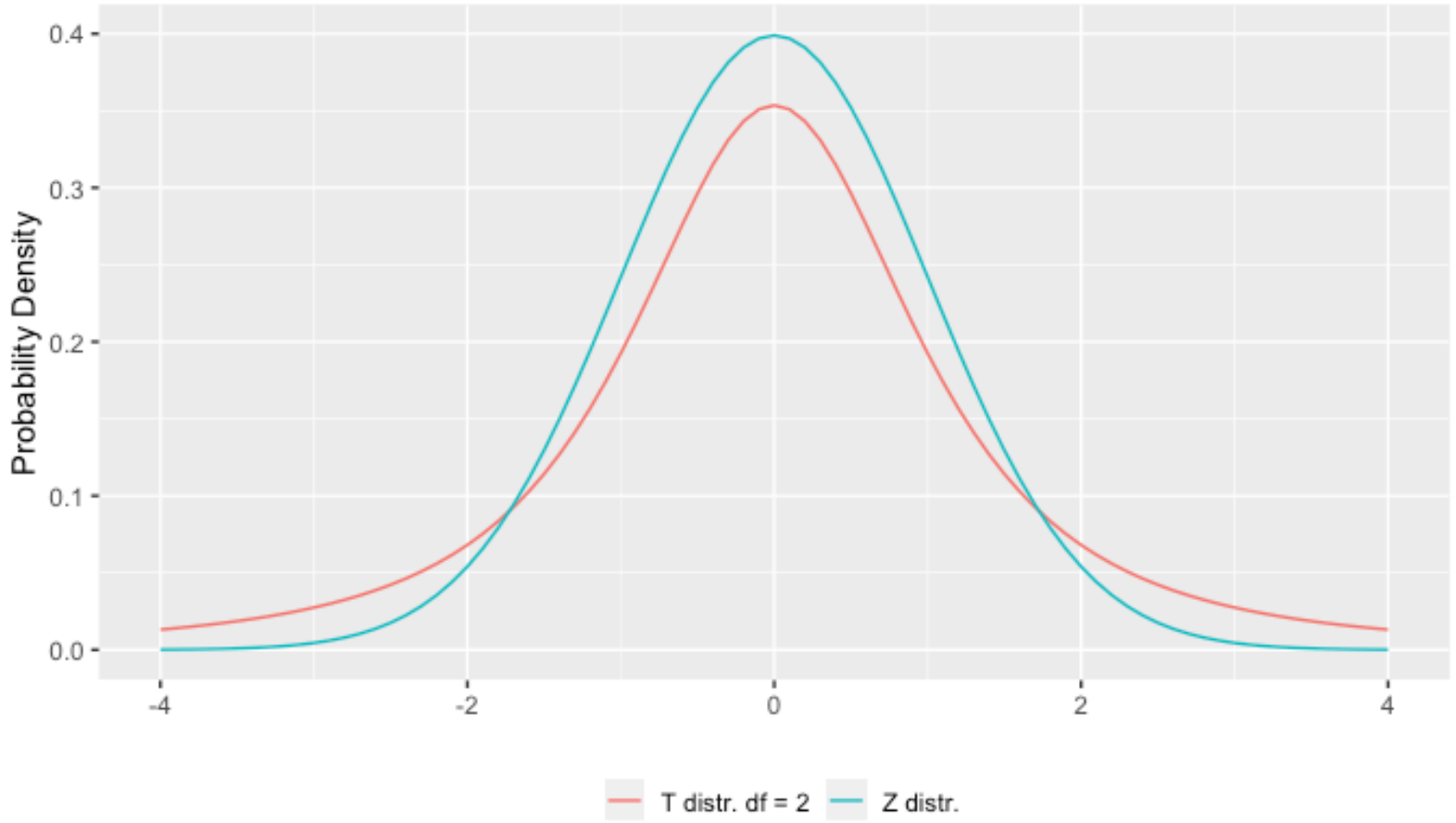
- T-distribution helps to account for additional uncertainty

How the T-distribution Works

- T-distribution helps to account for additional uncertainty when we **do not** know the population mean & standard deviation
  - More density at the extremes (“heavier tails”)

How the T-distribution Works

- T-distribution helps to account for additional uncertainty when we **do not** know the population mean & standard deviation
  - More density at the extremes (“heavier tails”)



Changes in Probability: T- vs. Z-distribution

https://rpsychologist.com/d3/tdist/

Changes in Probability: T- vs. Z-distribution

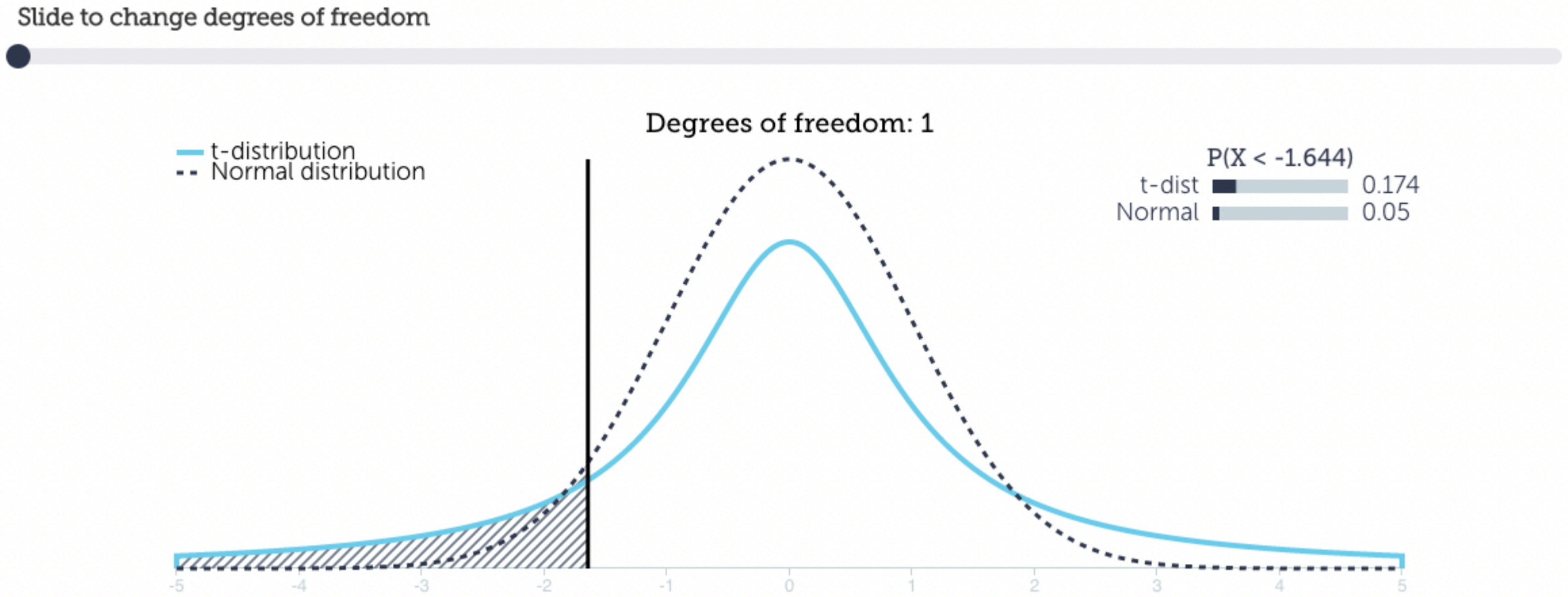
https://rpsychologist.com/d3/tdist/

Question: Do we have too little beer in our bottles?

Changes in Probability: T- vs. Z-distribution

https://rpsychologist.com/d3/tdist/

Question: Do we have too little beer in our bottles?



T-distribution changes shape with different degrees of freedom

- Here degrees of freedom (df) equal the sample size minus 1 (N – 1)

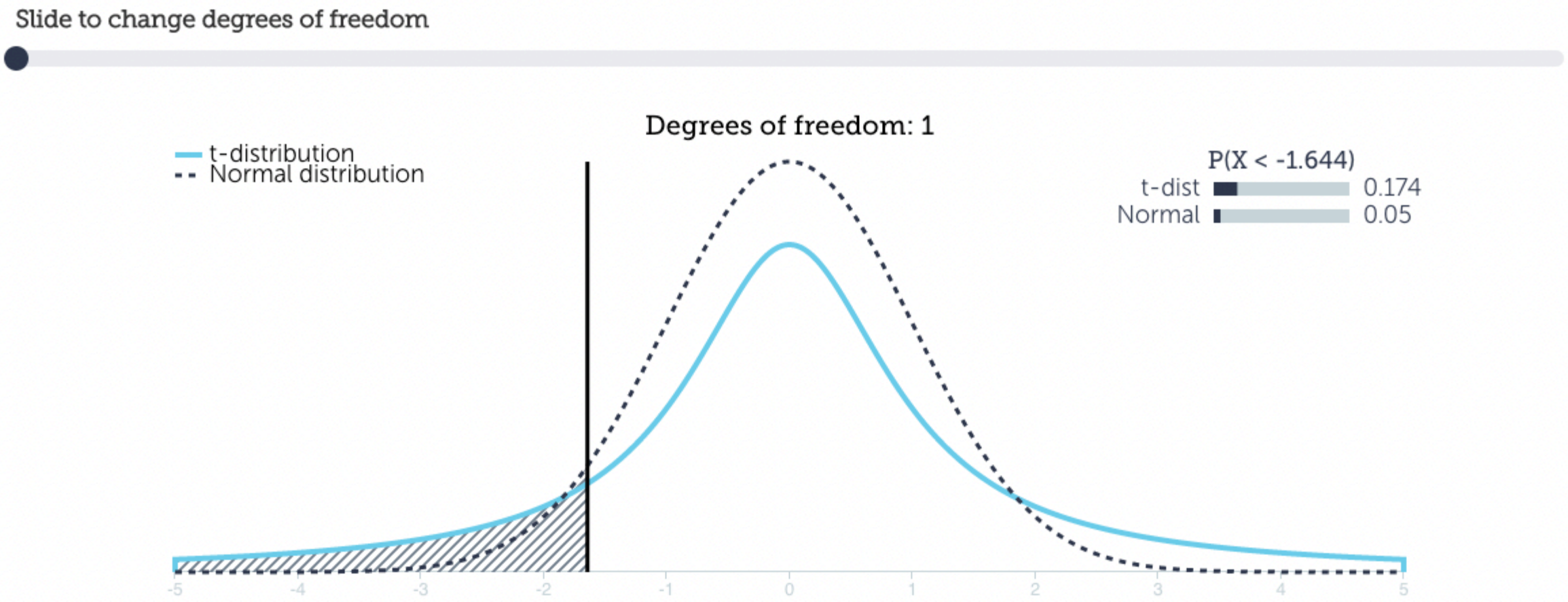
T-distribution changes shape with different degrees of freedom

- Here degrees of freedom (df) equal the sample size minus 1 (N – 1)
  - Just like the sample variance

T-distribution changes shape with different degrees of freedom

- Here degrees of freedom (df) equal the sample size minus 1 (N – 1)
  - Just like the sample variance

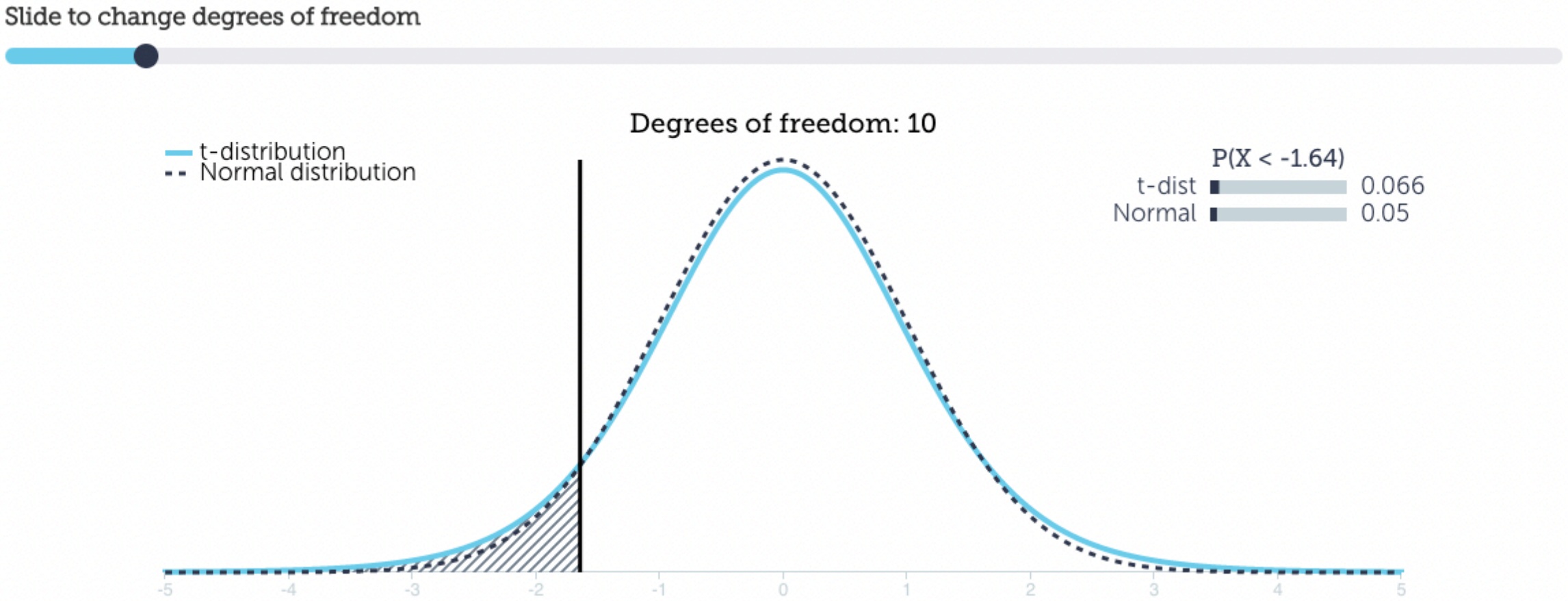
Question: Do we have too little beer in our bottles?



T-distribution changes shape with different degrees of freedom

- Here degrees of freedom (df) equal the sample size minus 1 (N – 1)
  - Just like the sample variance

Question: Do we have too little beer in our bottles?



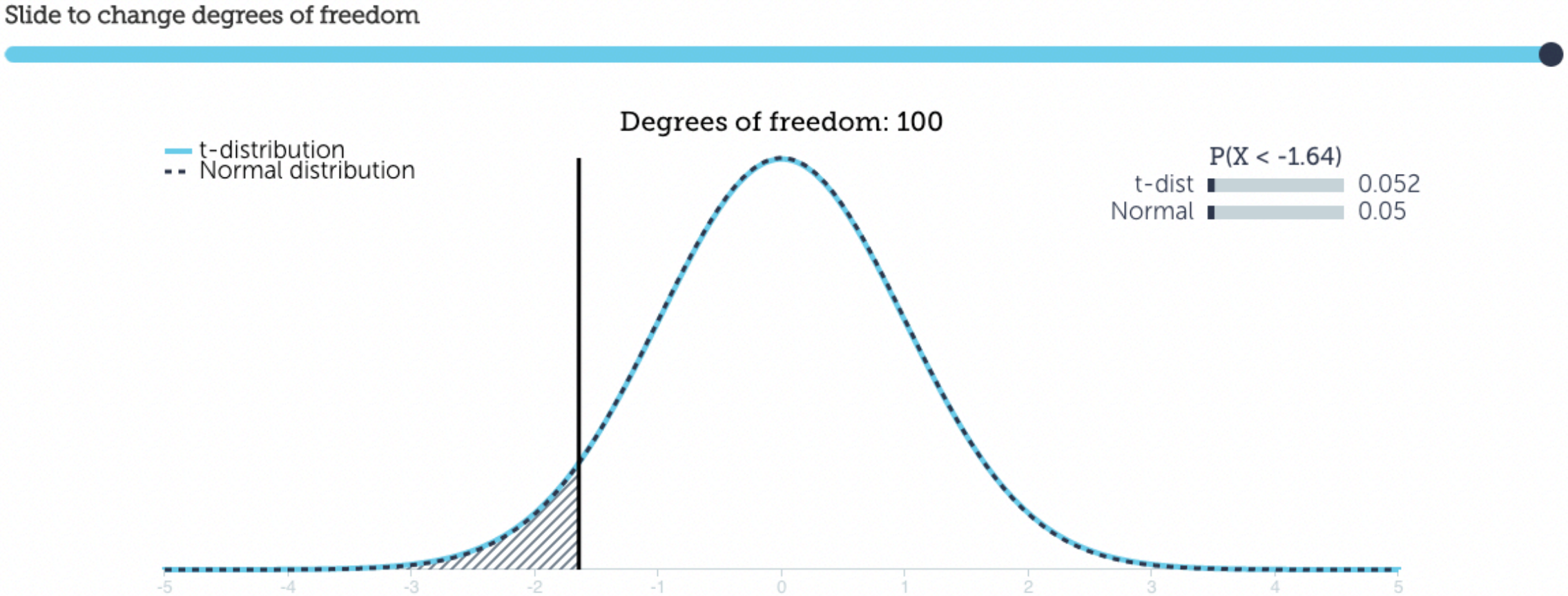
T-distribution changes shape with different degrees of freedom

- As degrees of freedom (sample size) increase, the t-distribution approaches the z-distribution

T-distribution changes shape with different degrees of freedom

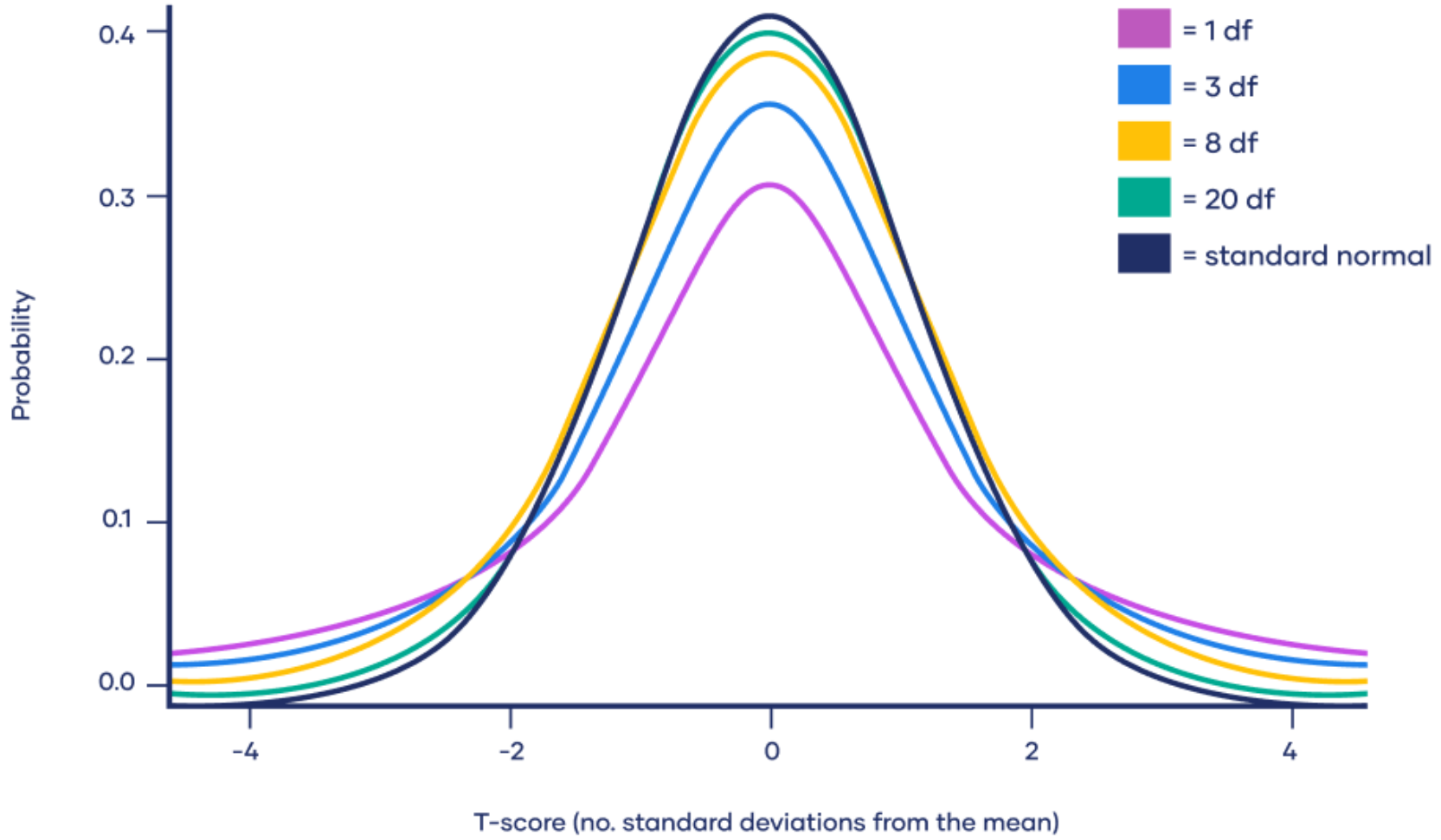
- As degrees of freedom (sample size) increase, the T-distribution approaches the Z-distribution

Question: Do we have too little beer in our bottles?



T-distribution becomes the Z-distribution when df goes to infinity

- As degrees of freedom (sample size) increase, the T-distribution approaches the Z-distribution



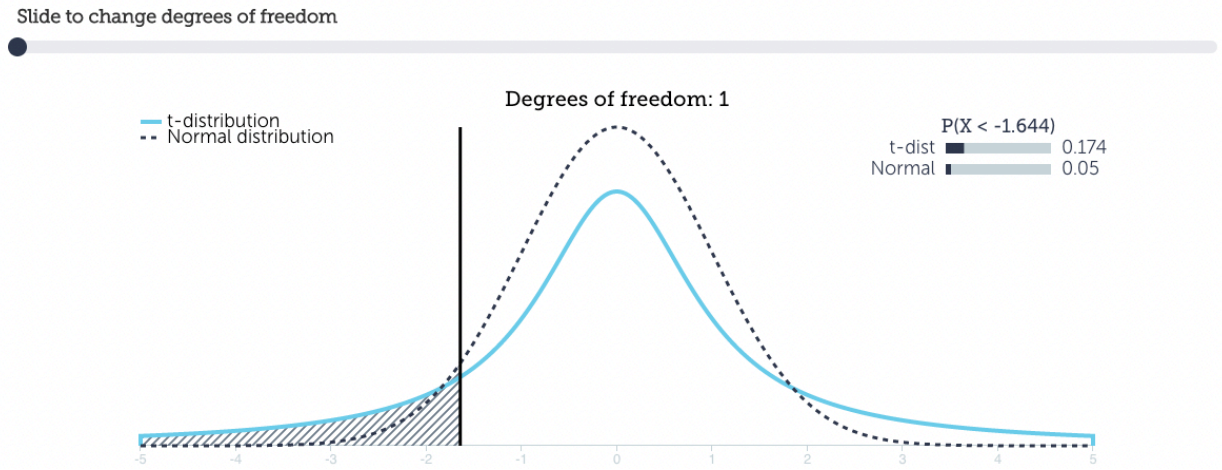
TLDR: T-Distribution

- When do not know the population variance and have small sample sizes, we need to adjust our expectations for sampling variation

TLDR: T-Distribution

- When do not know the population variance and have small sample sizes, we need to adjust our expectations for sampling variation

- At small degrees of freedom, the T-distribution has “thick tails” = more probability for extreme values

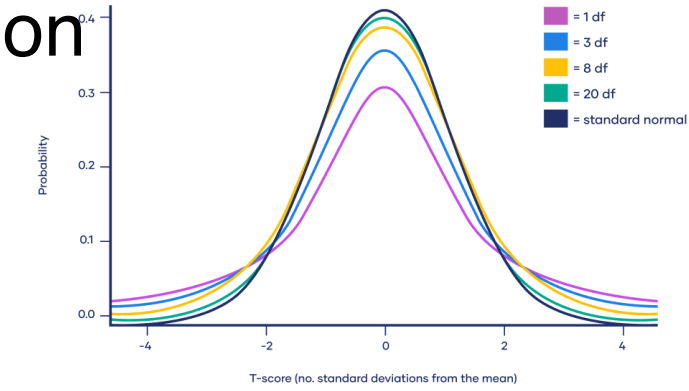


TLDR: T-Distribution

- When do not know the population variance and have small sample sizes, we need to adjust our expectations for sampling variation

- At small degrees of freedom, the T-distribution has “thick tails” = more probability for extreme values

- As sample size increases (large df’s), the T-distribution becomes more and more similar to the Z-distribution



Questions?

Ethan M. McCormick

ethan.mccormick@radboudumc.nl