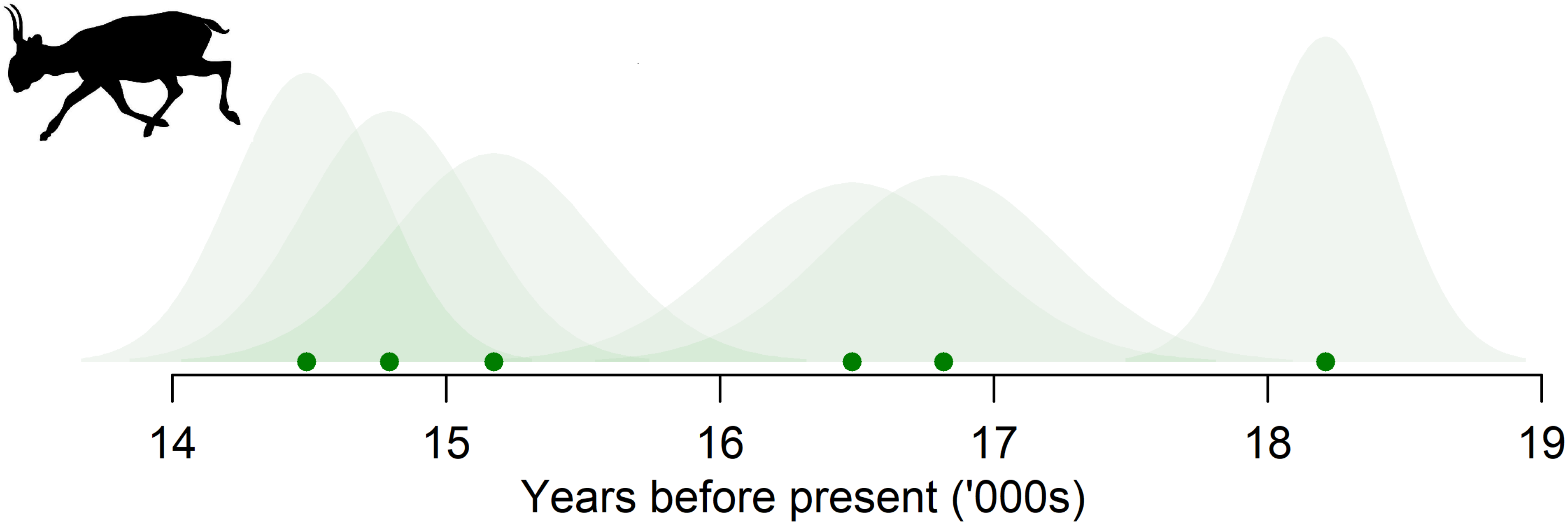
Estimating extinction time from the fossil record

using regression inversion

David Warton and Victor Tsang
School of Mathematics and Statistics and Evolution & Ecology Research Centre
University of New South Wales Sydney

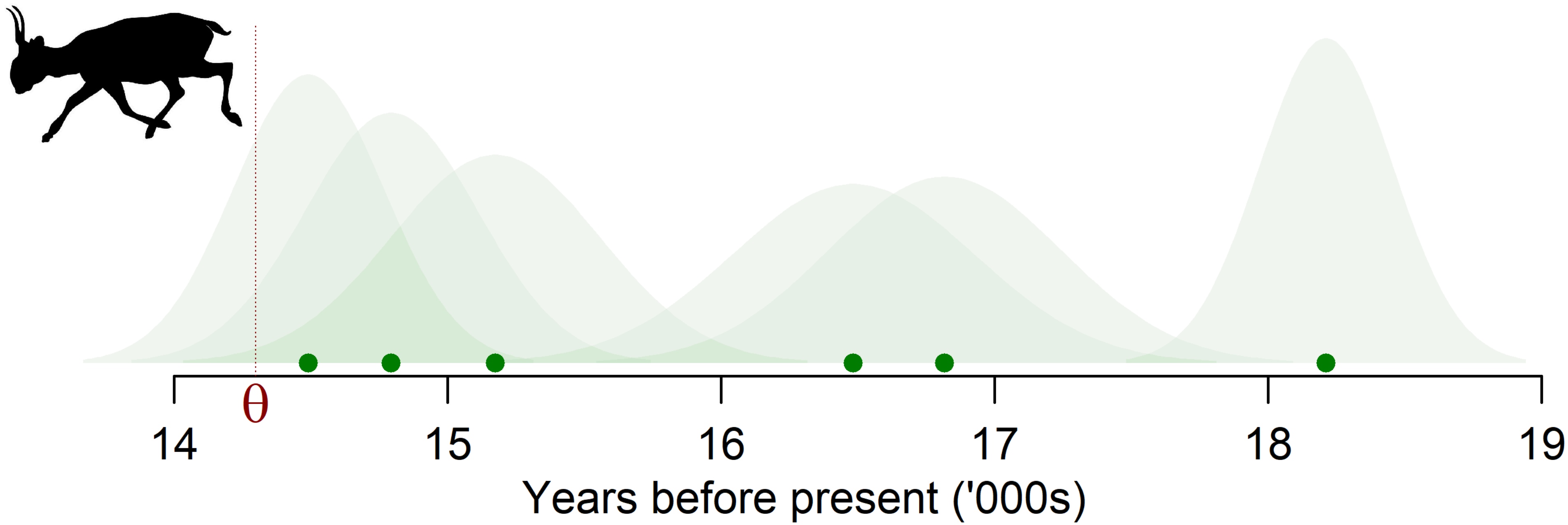
November 26th, 2025

Estimating extinction time from the fossil record



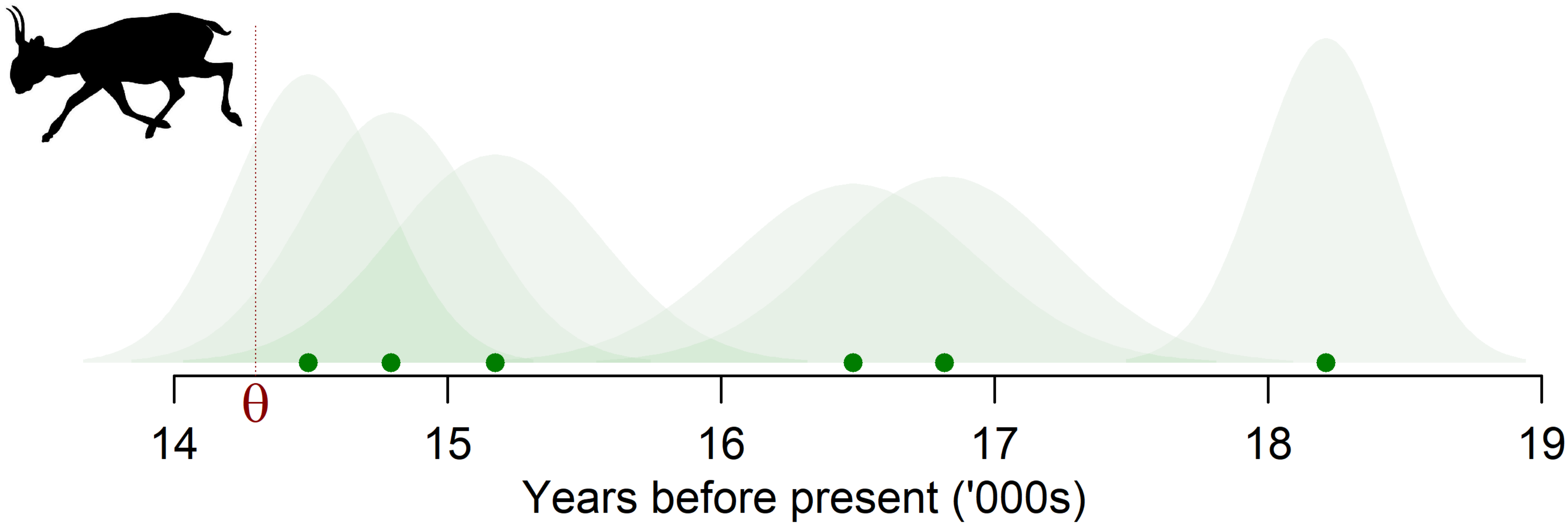
- We have a set of specimens from the fossil record with estimated ages
- Ages are measured with a known amount of error

Estimating extinction time from the fossil record



- We have a set of specimens from the fossil record with estimated ages
- Ages are measured with a known amount of error
- We want to estimate extinction time θ

Key data properties



We need to deal with

- *Sampling error* – last time it was seen isn't the last time it was there
- *Measurement error* – ages are estimated with error

Previous work

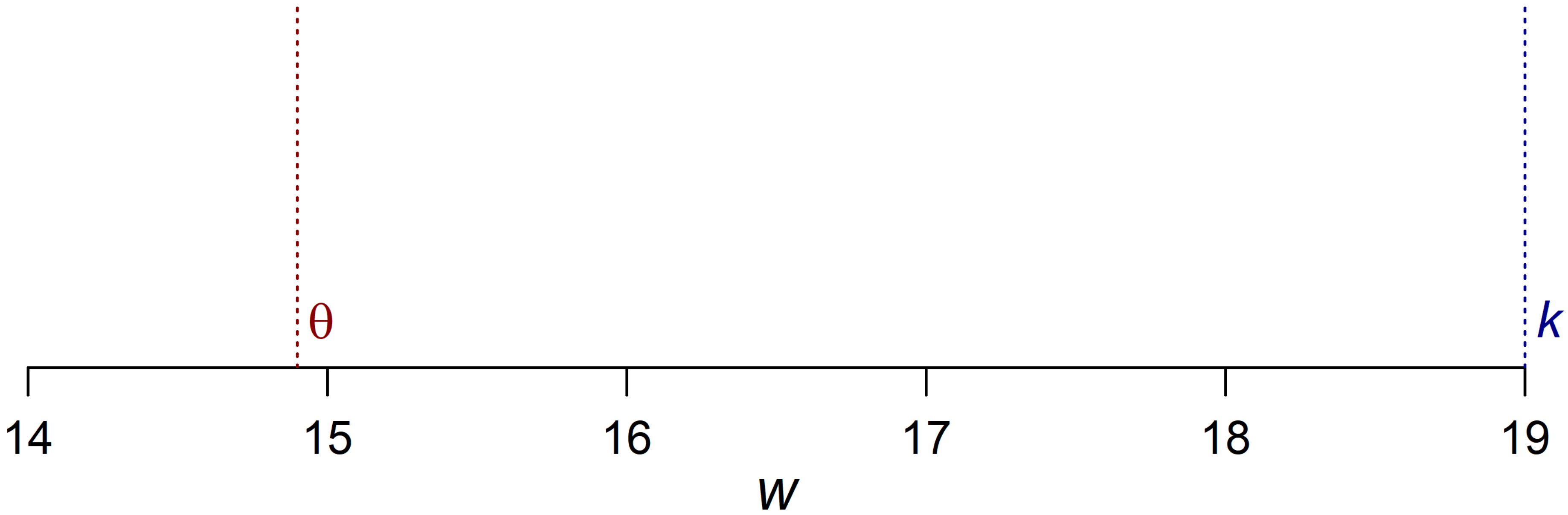
Who	Sampling error	Measurement error	Small samples
Strauss and Sadler (1989)	✓	×	✓
`GRIWM` by Bradshaw et al. (2012)	×	✓	✓
Solow, Roberts, and Robbirt (2006)	✓	✓	×

Aim

Who	Sampling error	Measurement error	Small samples
Strauss and Sadler (1989)	✓	×	✓
`GRIWM` by Bradshaw et al. (2012)	×	✓	✓
Solow, Roberts, and Robbirt (2006)	✓	✓	×
This talk	**✓**	**✓**	**✓**

The Compound Uniform-Truncated T Distribution

We observe $W = X + \epsilon$ which can be no larger than k.

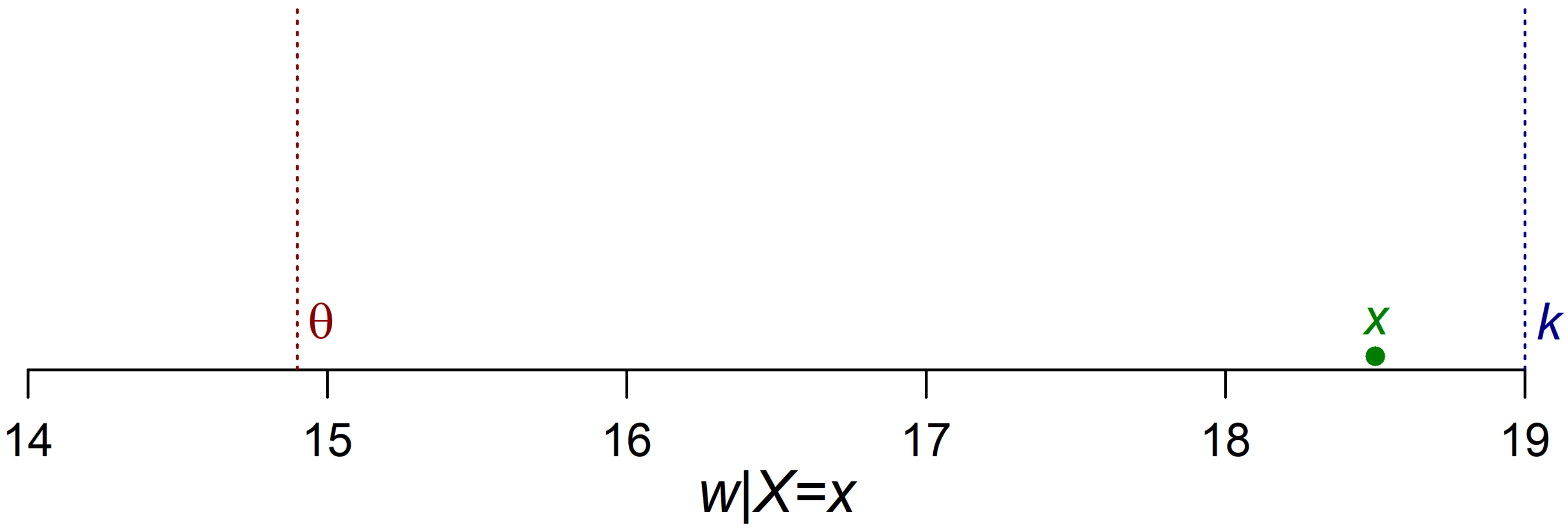


X is the true fossil age, which can be no smaller than θ.

ϵ is measurement error with known standard deviation σ.

The Compound Uniform-Truncated T Distribution

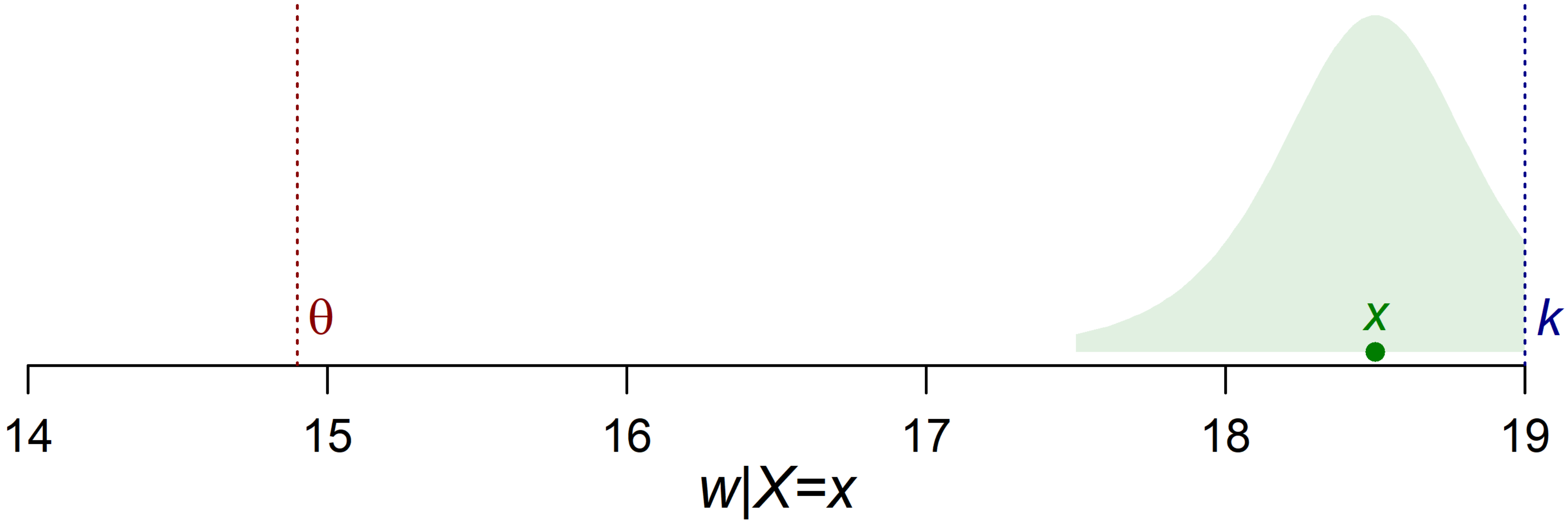
We propose defining a distribution for W via the conditionals of (X, ϵ):



- Conditional on $X = x \ldots$

The Compound Uniform-Truncated T Distribution

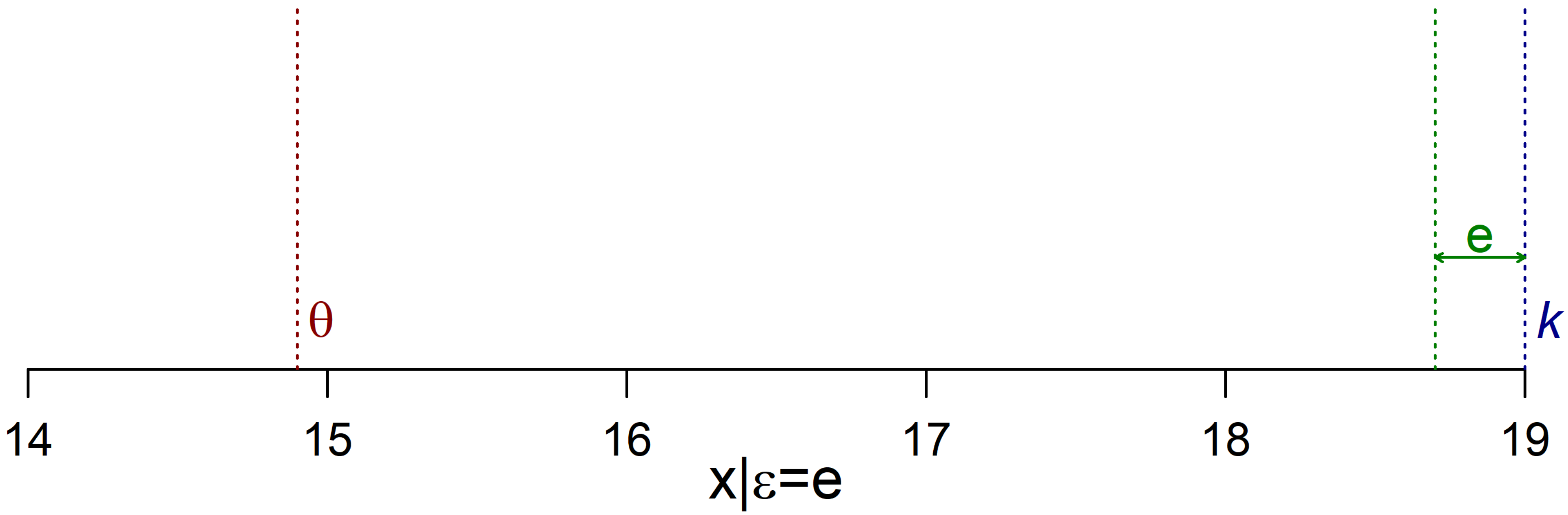
We propose defining a distribution for W via the conditionals of (X, ϵ):



- Conditional on $X = x$, we assume $\epsilon | X \stackrel{d}{=} (T | T < k - x)$, where $T \sim t_\nu$

The Compound Uniform-Truncated T Distribution

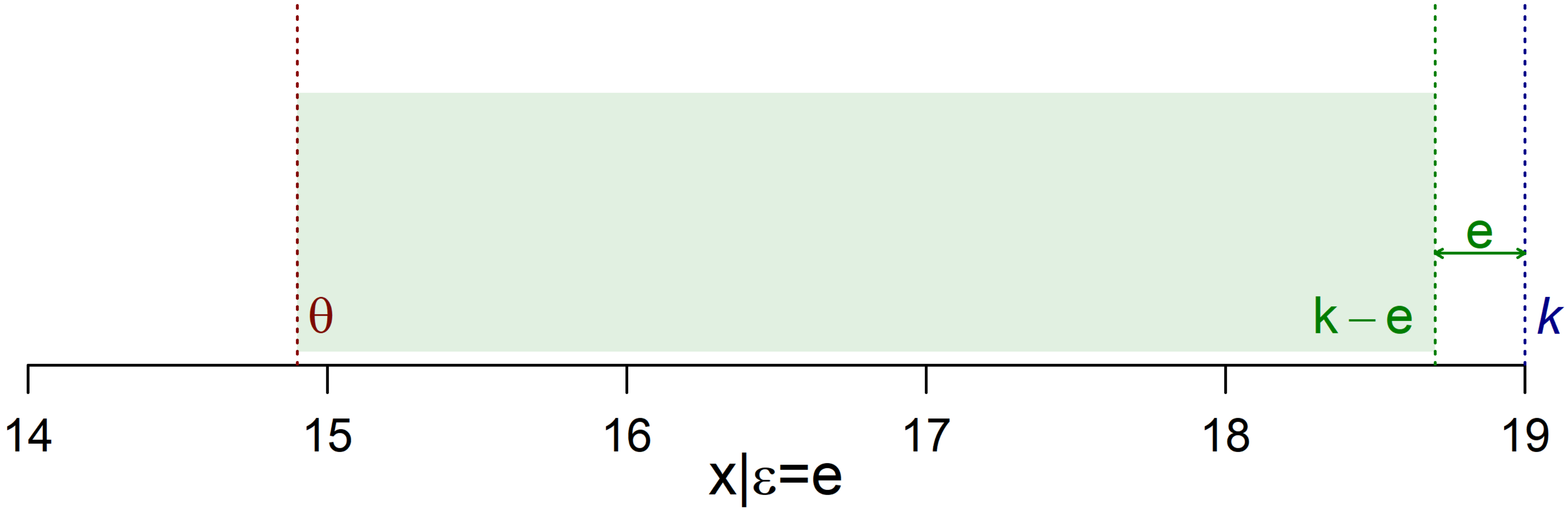
We propose defining a distribution for W via the conditionals of (X, ϵ):



- Conditional on $X = x$, we assume $\epsilon | X \stackrel{d}{=} (T | T < k - x)$, where $T \sim t_\nu$
- Conditional on $\epsilon = e \ldots$

The Compound Uniform-Truncated T Distribution

We propose defining a distribution for W via the conditionals of (X, ϵ):



- Conditional on $X = x$, we assume $\epsilon | X \stackrel{d}{=} (T | T < k - x)$, where $T \sim t_\nu$
- Conditional on $\epsilon = e$, we assume $X | \epsilon \sim \mathcal{U}[\theta, k - e]$

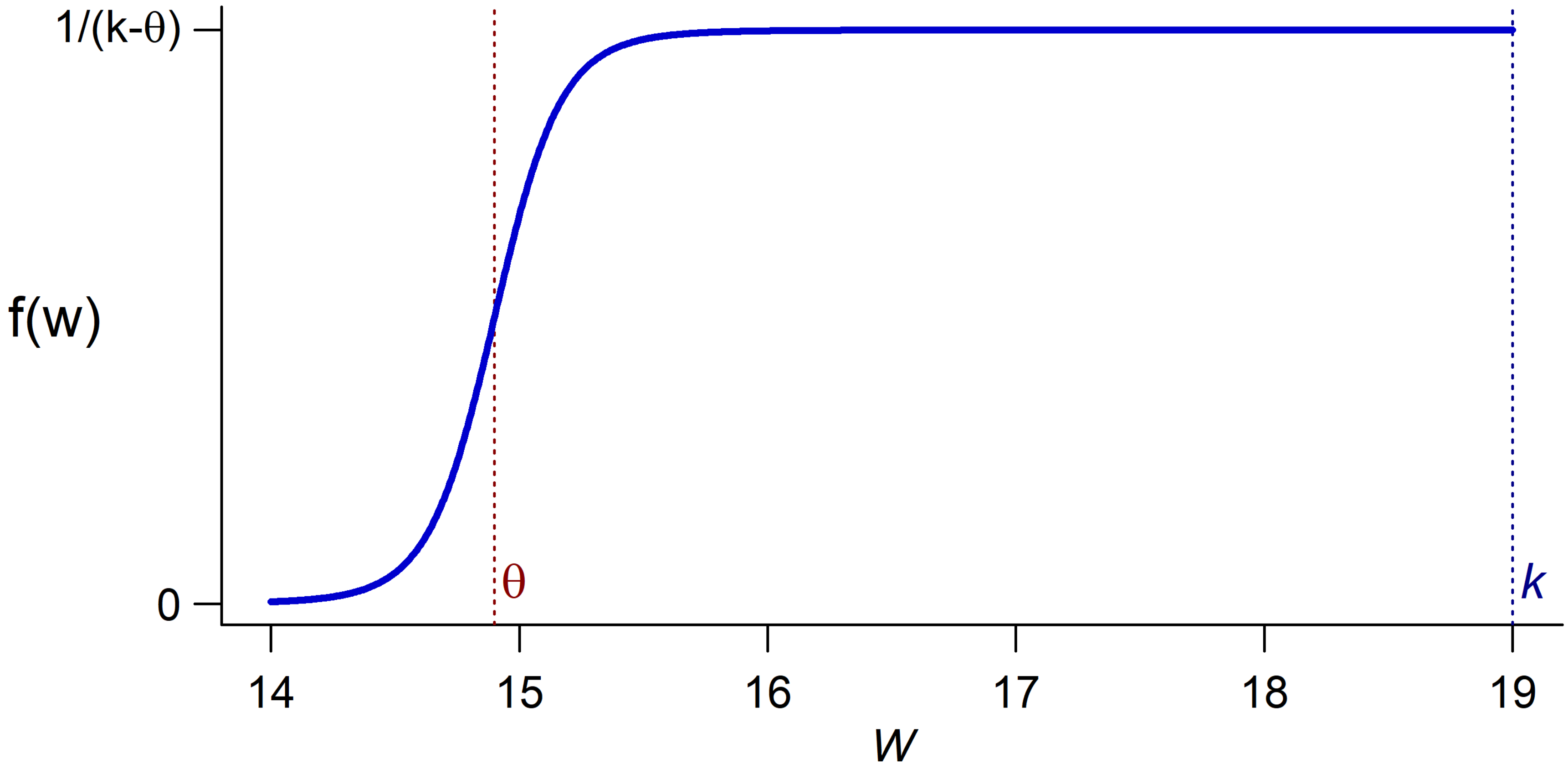
The Compound Uniform-Truncated T Distribution

We can then show that the probability density function of W is:

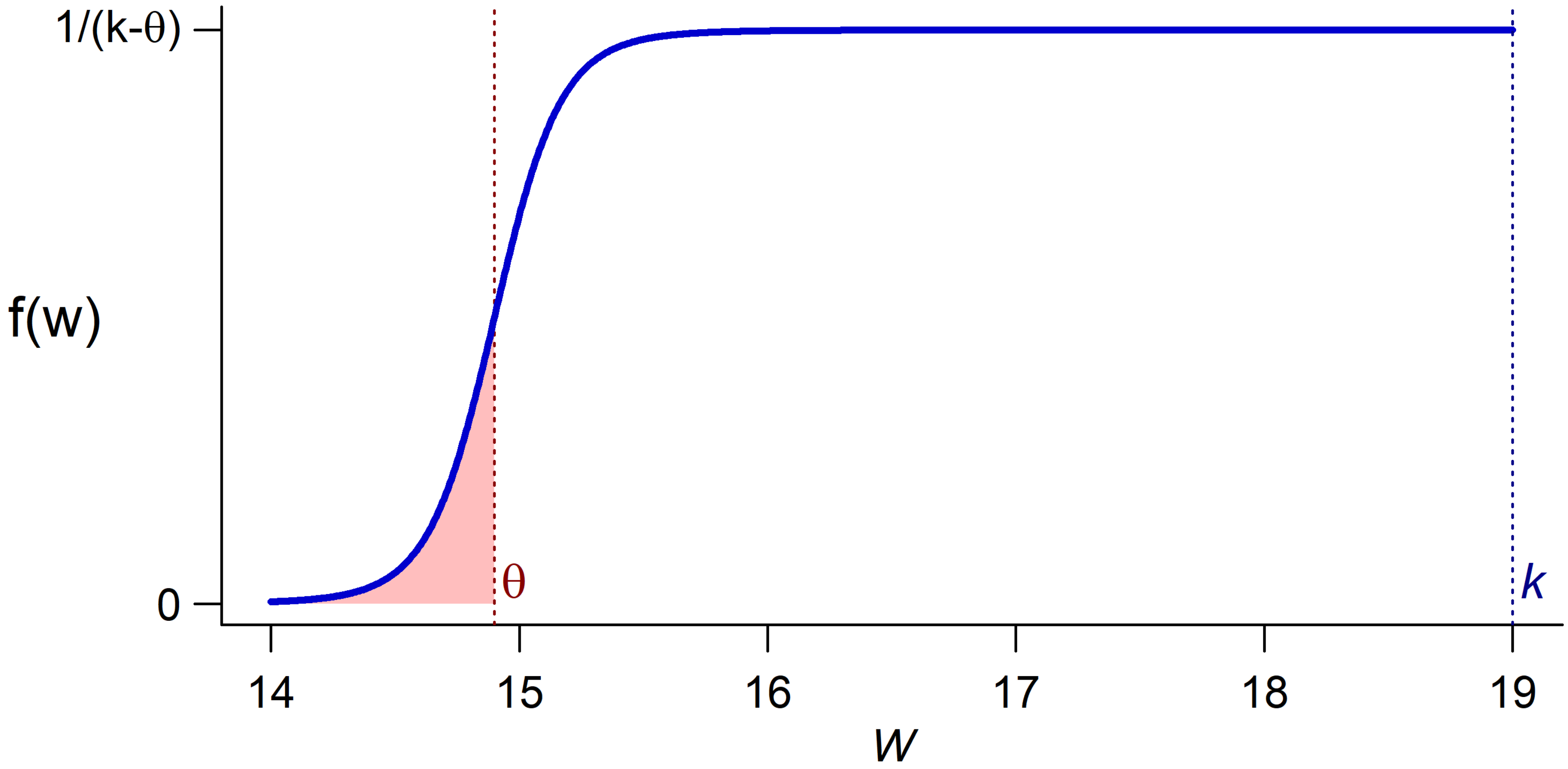
$$f(w) = \frac{G_\nu\left(\frac{w-\theta}{\sigma}\right)}{(k-\theta)G_\nu\left(\frac{k-\theta}{\sigma}\right) + \sigma g_\nu\left(\frac{k-\theta}{\sigma}\right)\left(\frac{\nu}{\nu-1}\right)\left(1+\frac{(k-\theta)^2}{\sigma^2\nu}\right)}$$

where $G_\nu(\cdot)$ and $g_\nu(\cdot)$ are the cumulative distribution function and the probability density function of Student's t_ν distribution.

The Compound Uniform-Truncated T Distribution

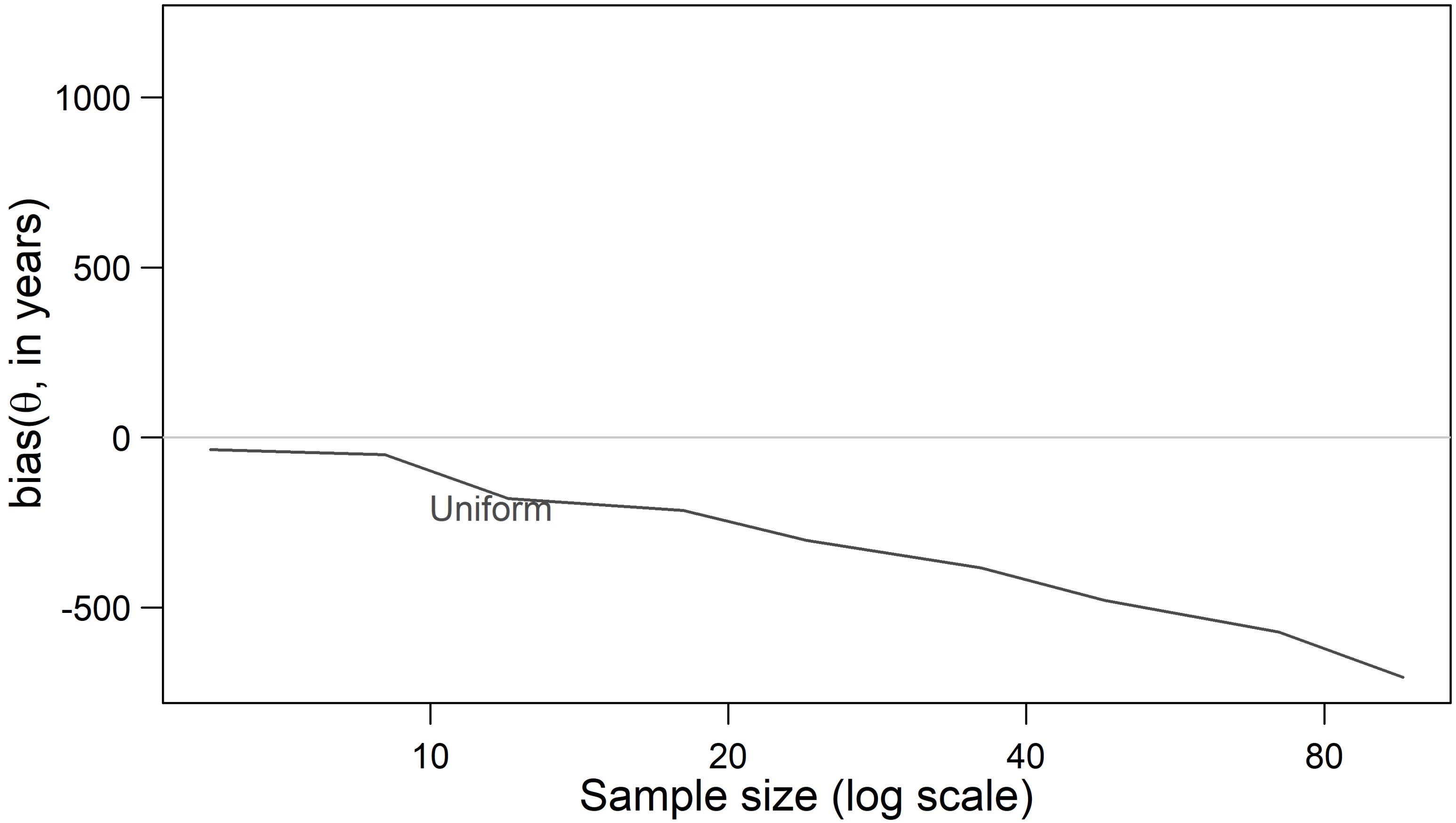


The Compound Uniform-Truncated T Distribution

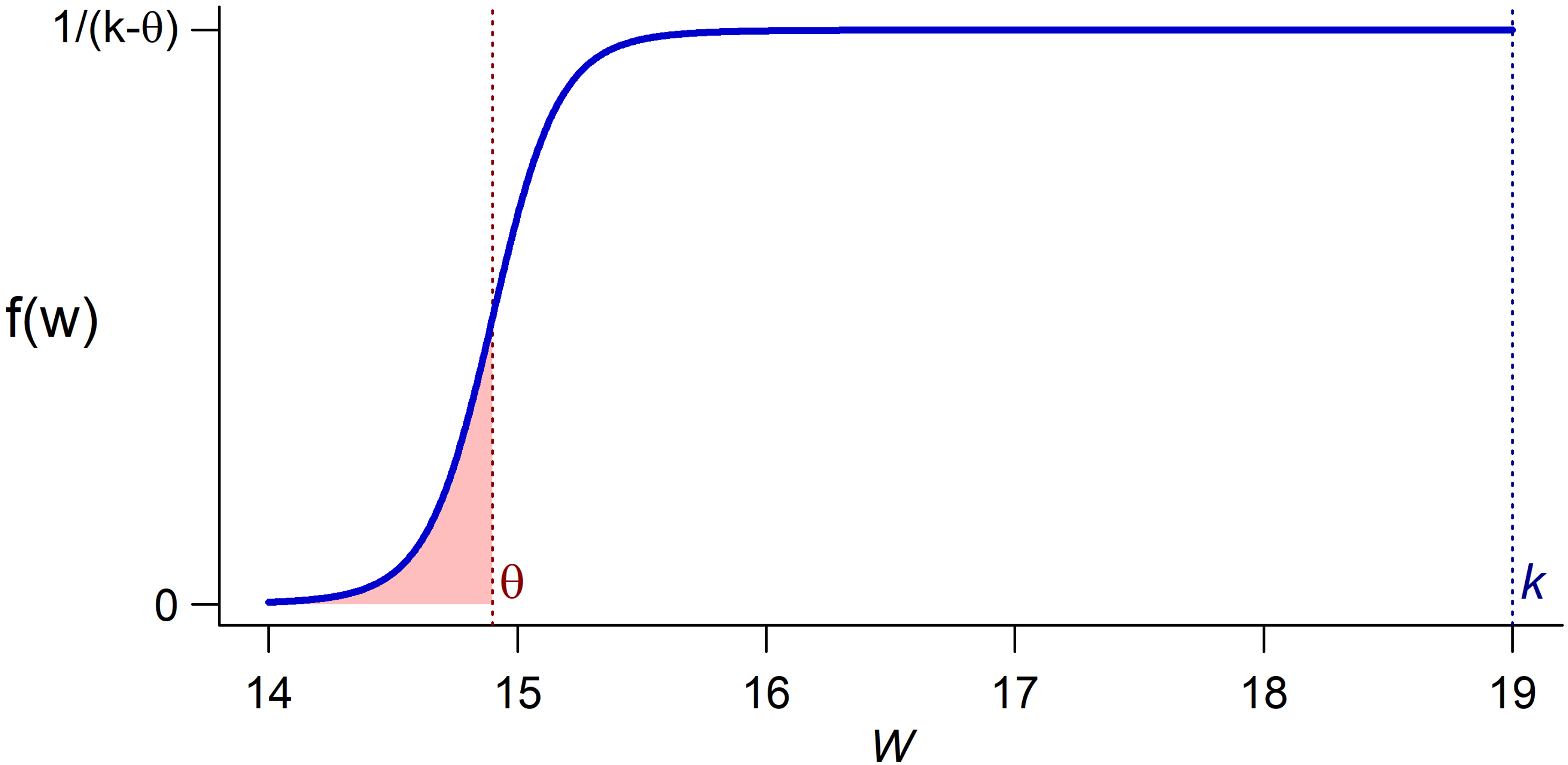


Some observed ages will be *younger* than true time of extinction!

Ignore measurement error $\rightarrow$ negative bias



The Compound Uniform-Truncated T Distribution



Some observed ages will be *younger* than true time of extinction!

Asymptotic Inference about θ?

1. Wald CI:

$$\left[\hat{\theta} - \frac{I(\hat{\theta})^{-1}}{\sqrt{n}}, \hat{\theta} + \frac{I(\hat{\theta})^{-1}}{\sqrt{n}}\right]$$

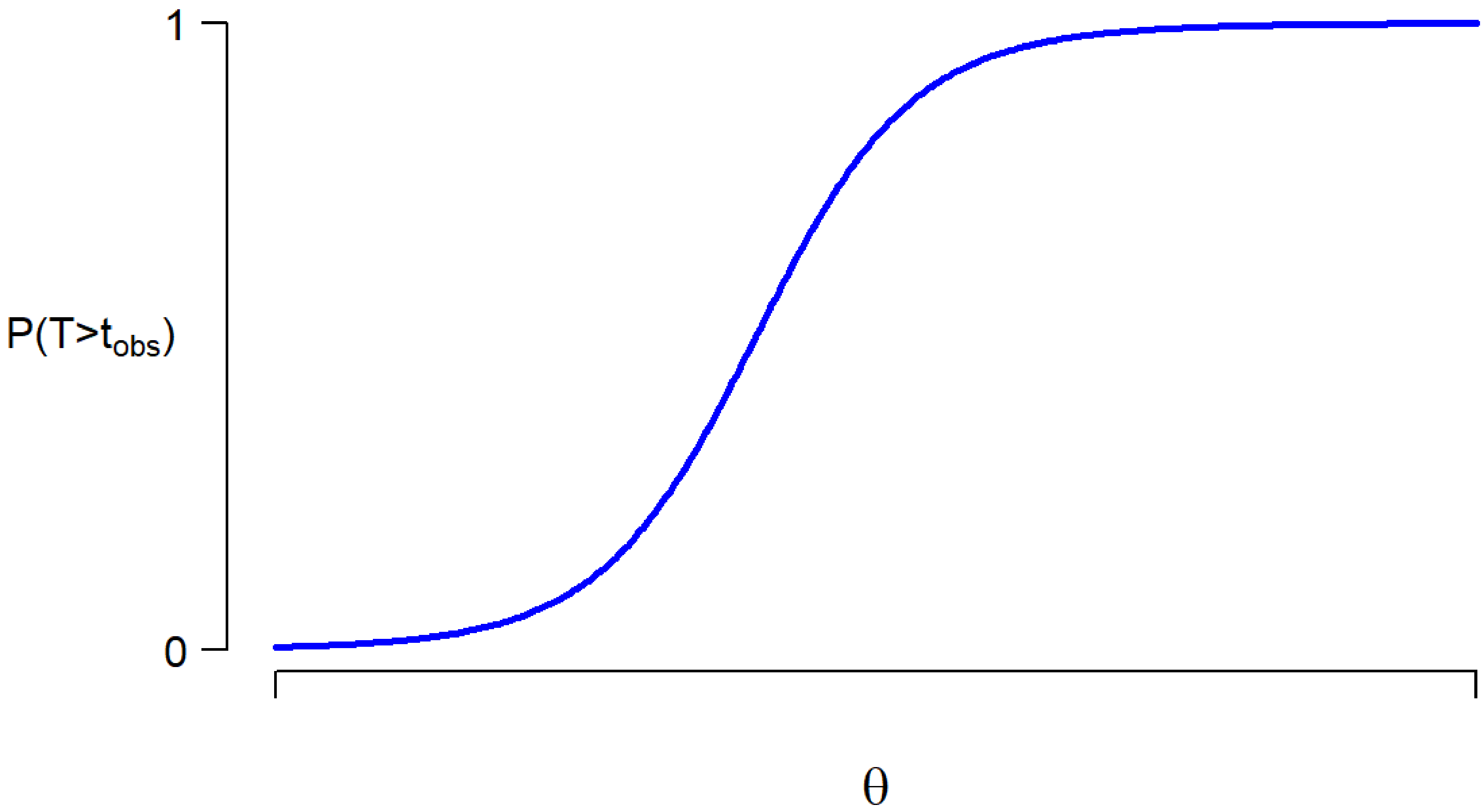
   since $\sqrt{n}I(\theta_0)(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0,1)$

2. Likelihood ratio test inversion:

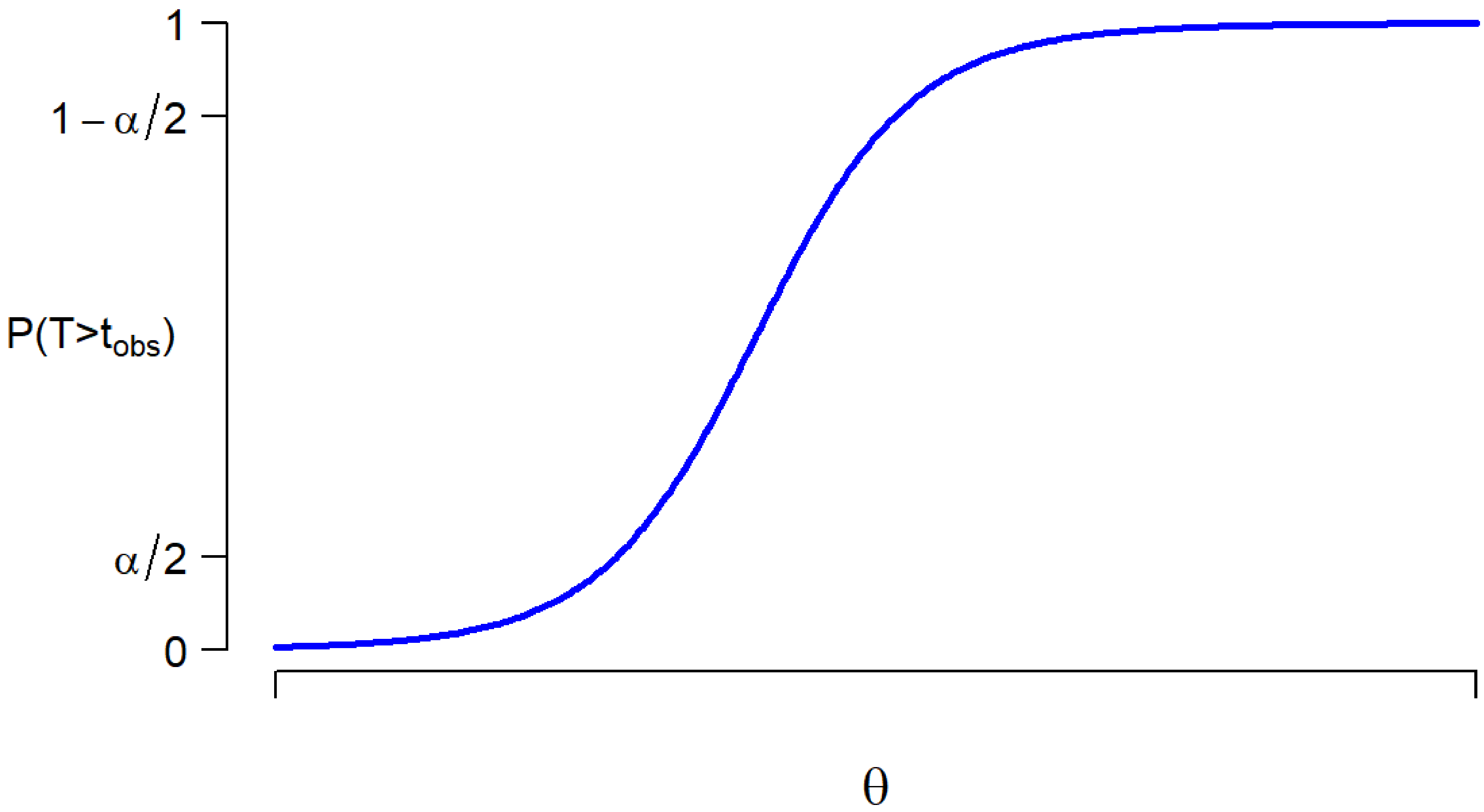
$$\{\theta : -2\log\Lambda(\theta) < q_\alpha\}$$

   where $\mathbb{P}(X^2 > q_\alpha) = \alpha$ for $X^2 \sim \chi^2_1$, since
   $-2\log\Lambda(\theta_0) = 2\ell(\hat{\theta}) - 2\ell(\theta_0) \xrightarrow{d} \chi^2_1$

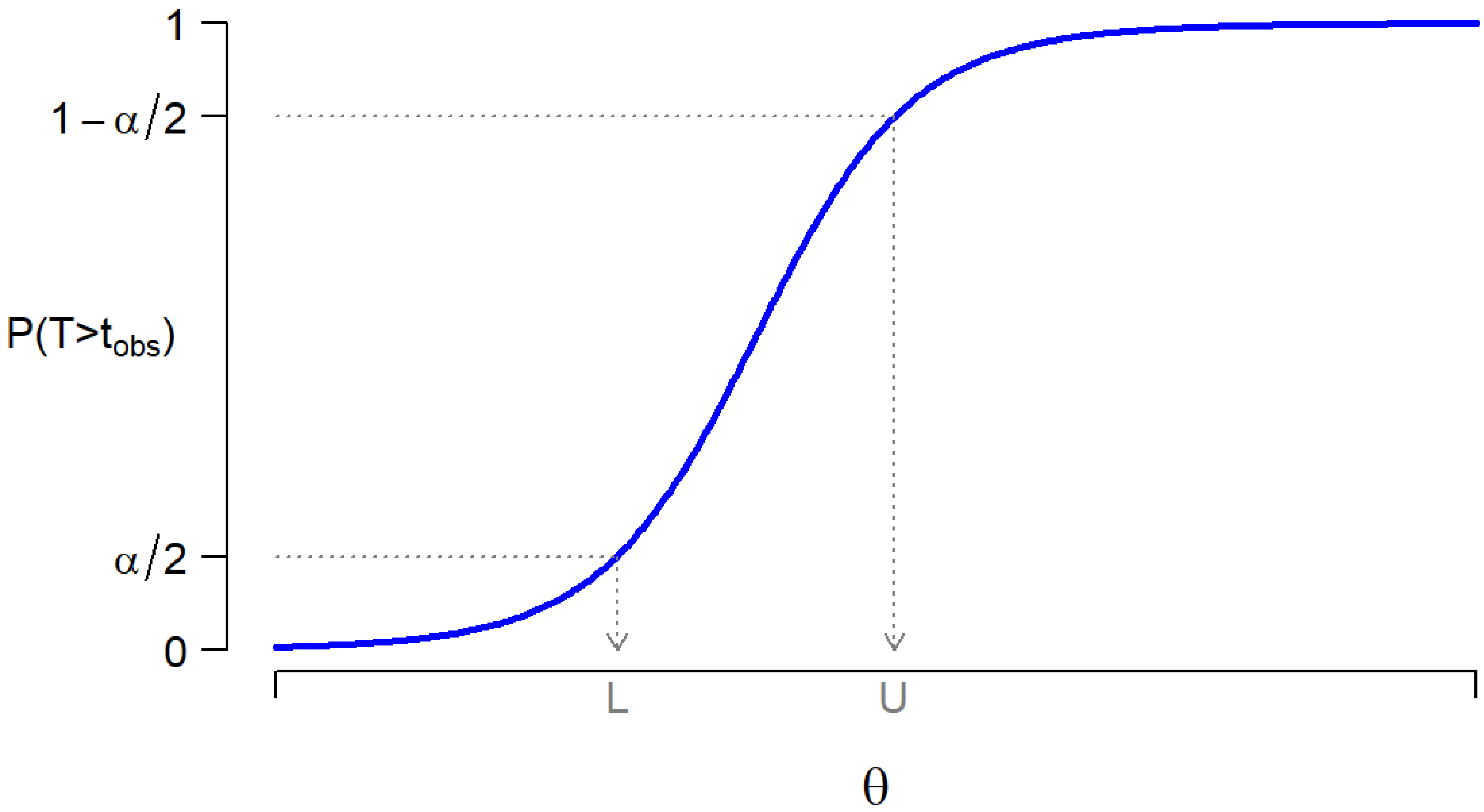
Confidence intervals by inversion



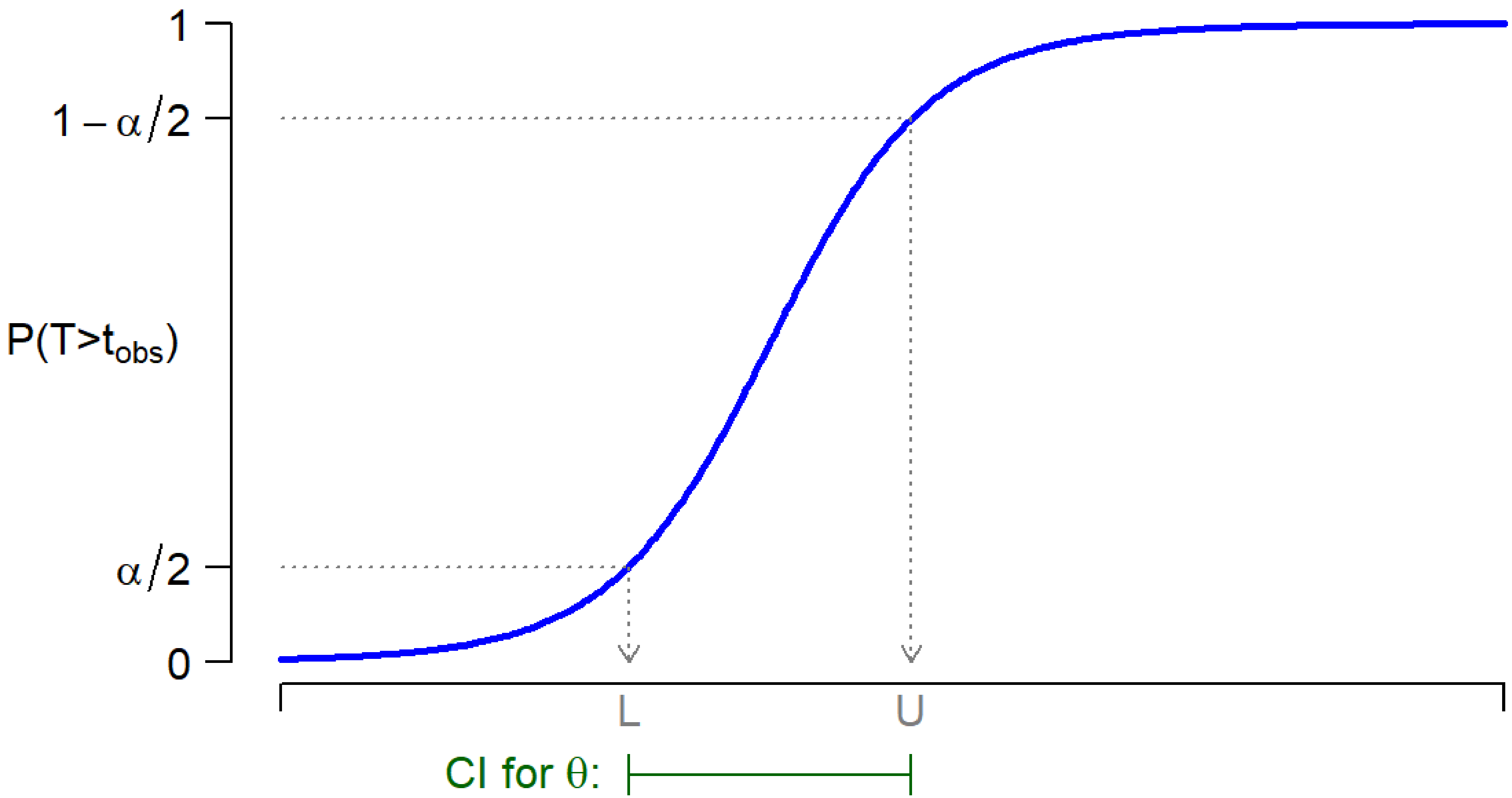
Confidence intervals by inversion



Confidence intervals by inversion



Confidence intervals by inversion



Asymptotic Inference about θ?

1. Wald CI:

$$\left[\hat{\theta} - \frac{I(\hat{\theta})^{-1}}{\sqrt{n}}, \hat{\theta} + \frac{I(\hat{\theta})^{-1}}{\sqrt{n}}\right]$$

   since $\sqrt{n} I(\theta_0)(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, 1)$

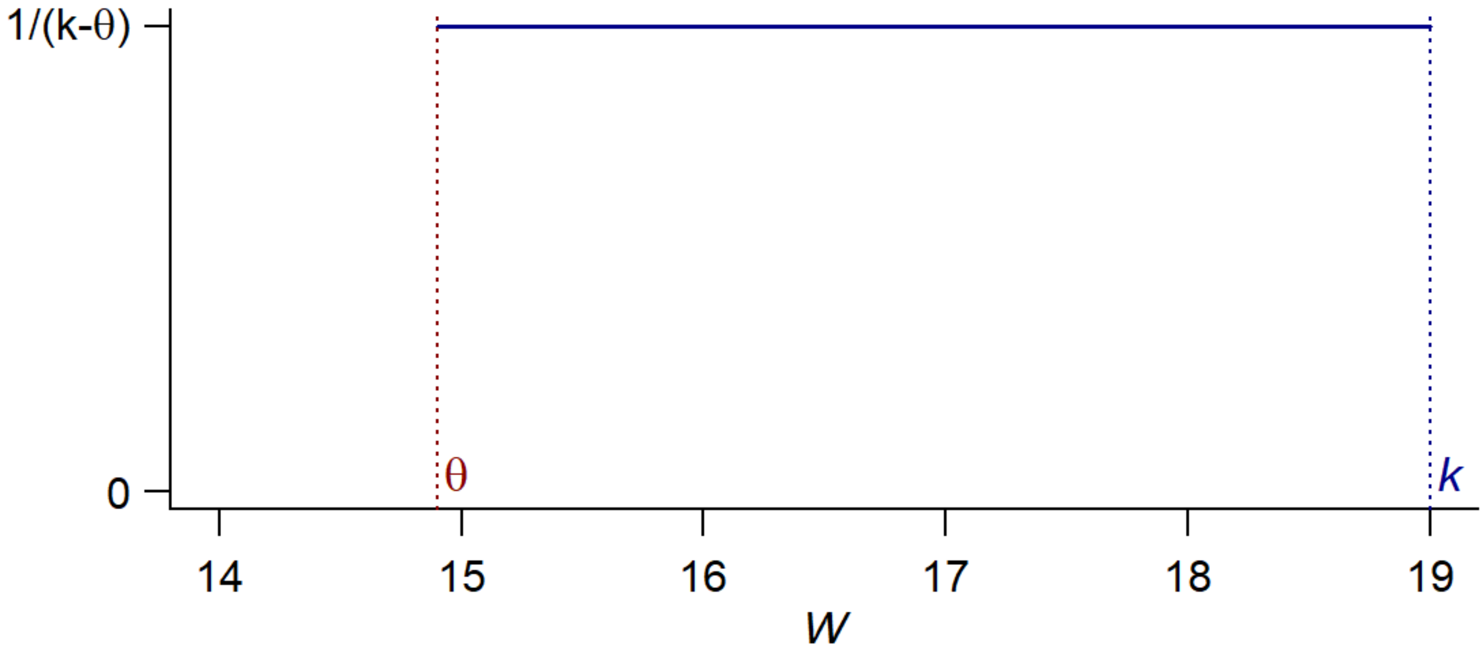
2. Likelihood ratio test inversion:

$$\{\theta : -2 \log \Lambda(\theta) < q_\alpha\}$$

   where $\mathbb{P}(X^2 > q_\alpha) = \alpha$ for $X^2 \sim \chi^2_1$, since
   $-2 \log \Lambda(\theta_0) = 2\ell(\hat{\theta}) - 2\ell(\theta_0) \xrightarrow{d} \chi^2_1$

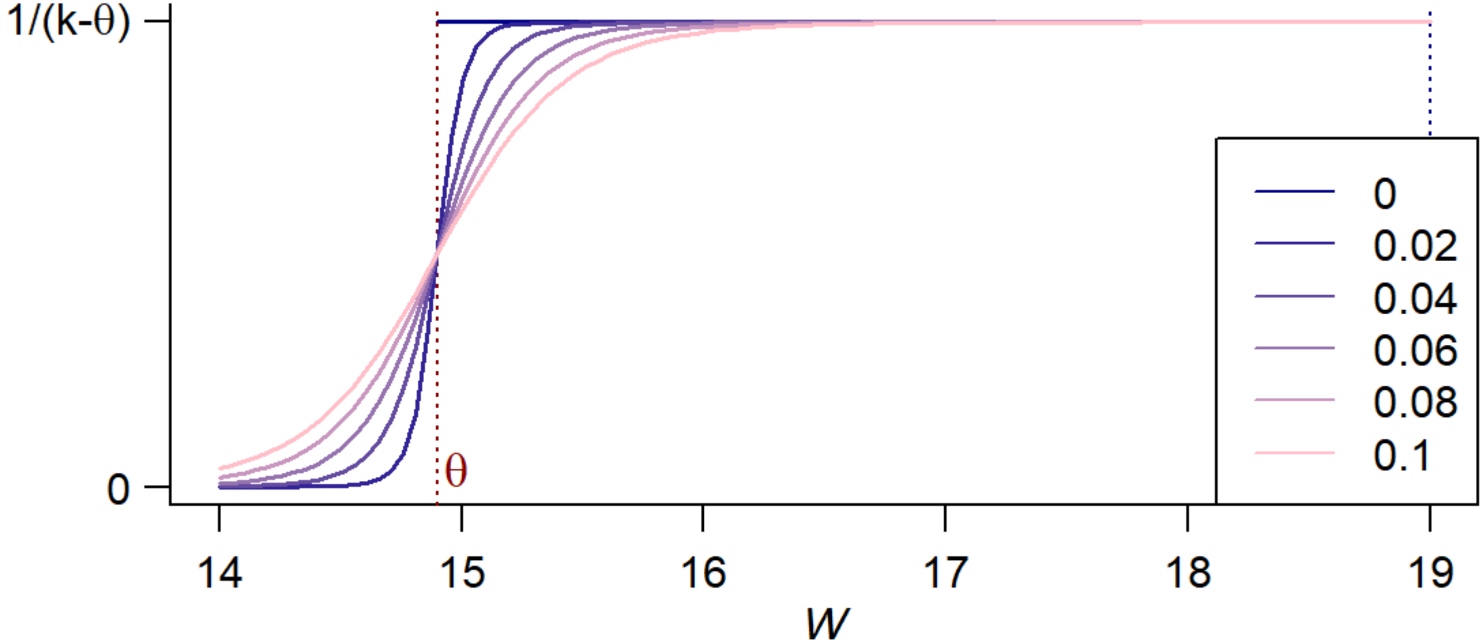
Problem: doesn't work for small σ (or small n)

We are in the neighbourhood of $\sigma = 0$



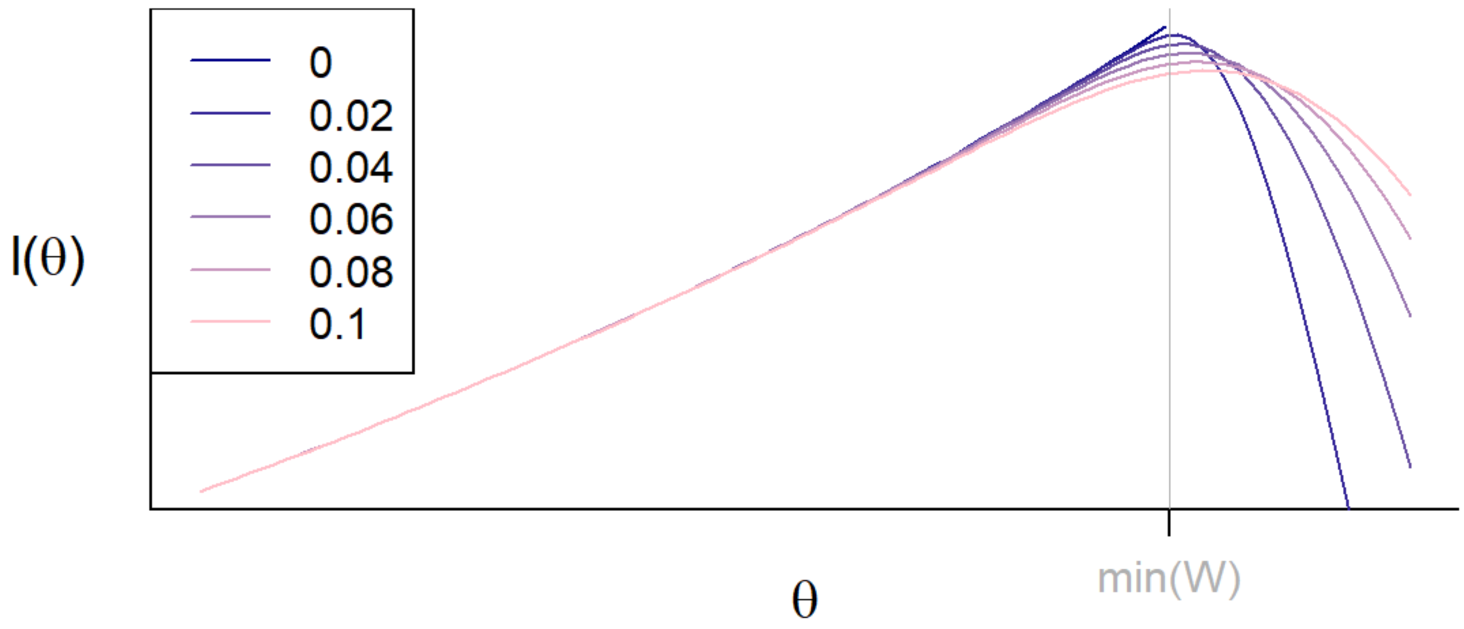
Problem: doesn't work for small σ (or small n)

We are in the neighbourhood of $\sigma = 0$



Problem: doesn't work for small σ (or small n)

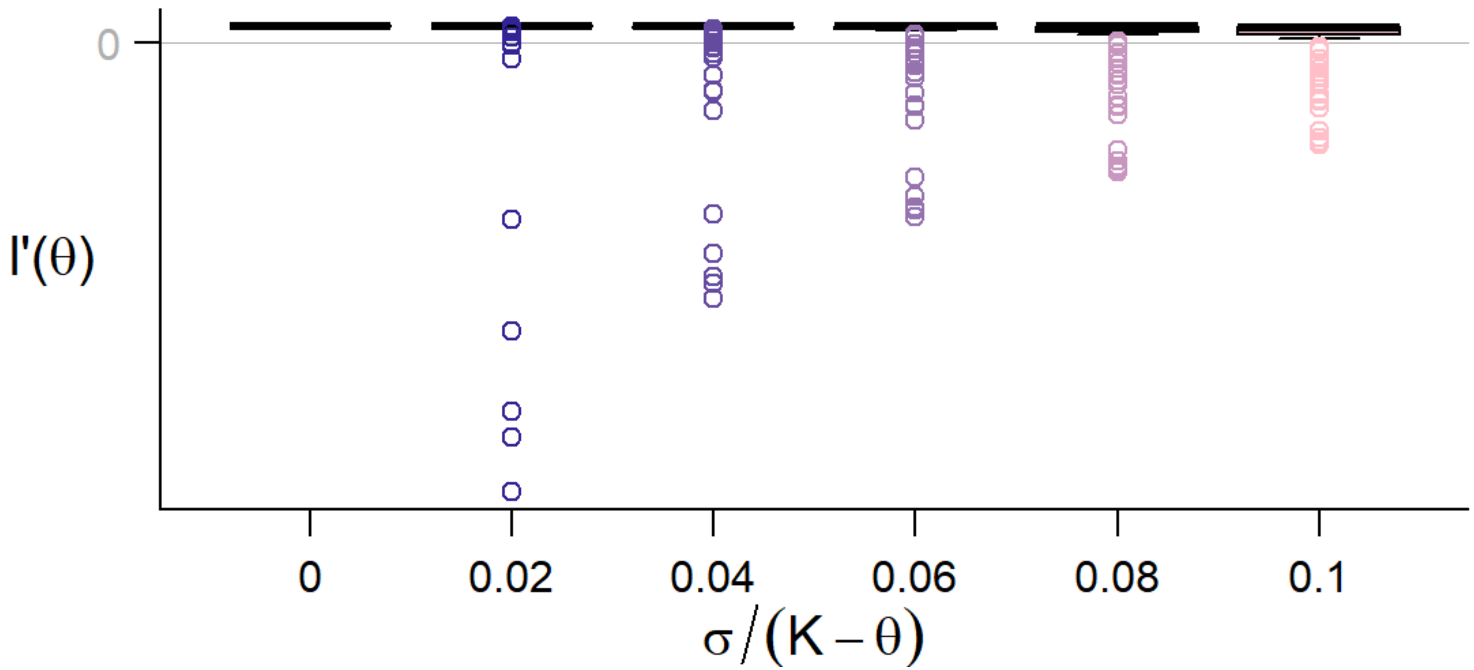
We are in the neighbourhood of $\sigma = 0$, where:



- the likelihood is not smooth

Problem: doesn't work for small σ (or small n)

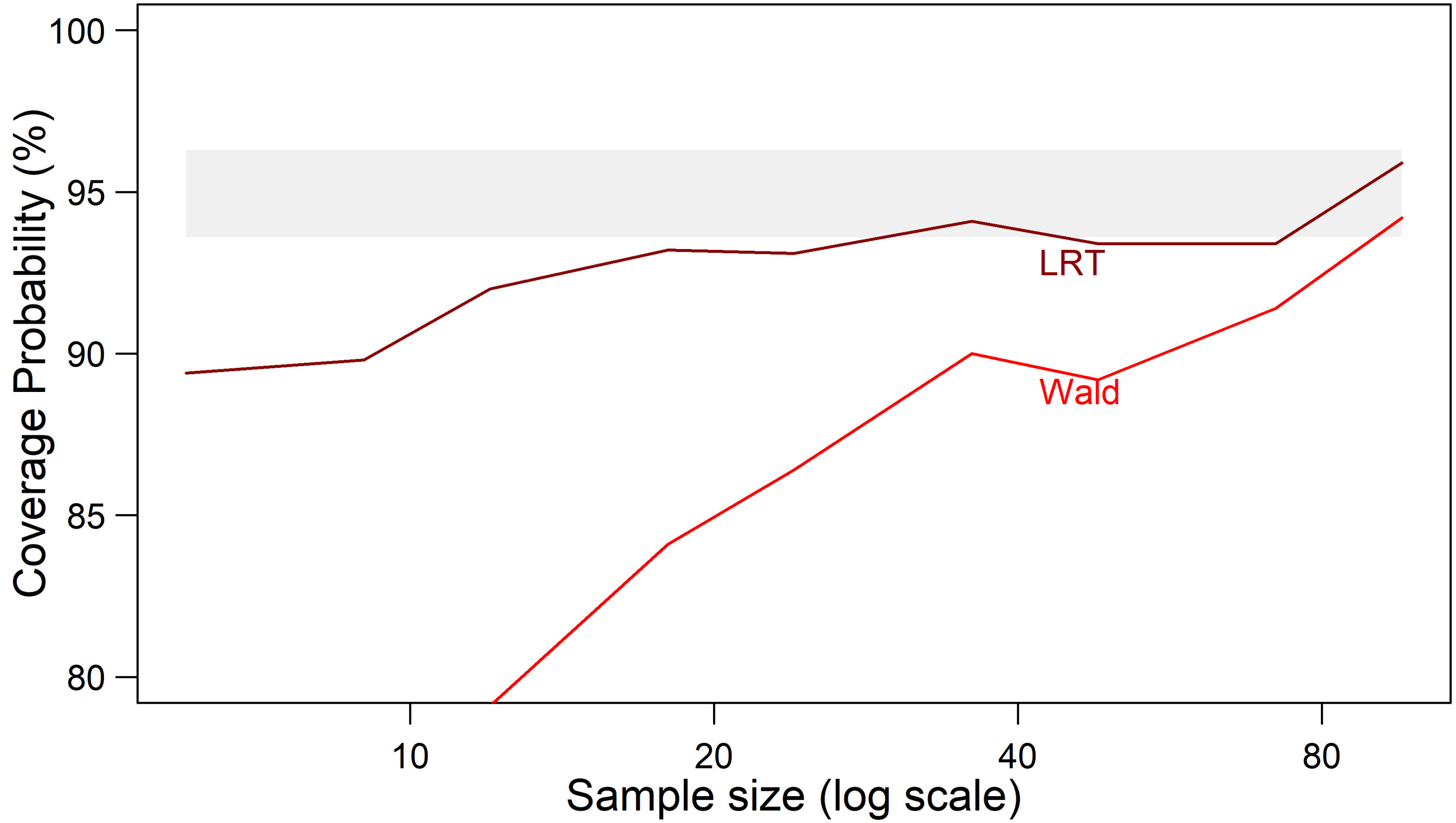
We are in the neighbourhood of $\sigma = 0$, where:



- the likelihood is not smooth
- Central Limit Theorem does not apply to the score equation

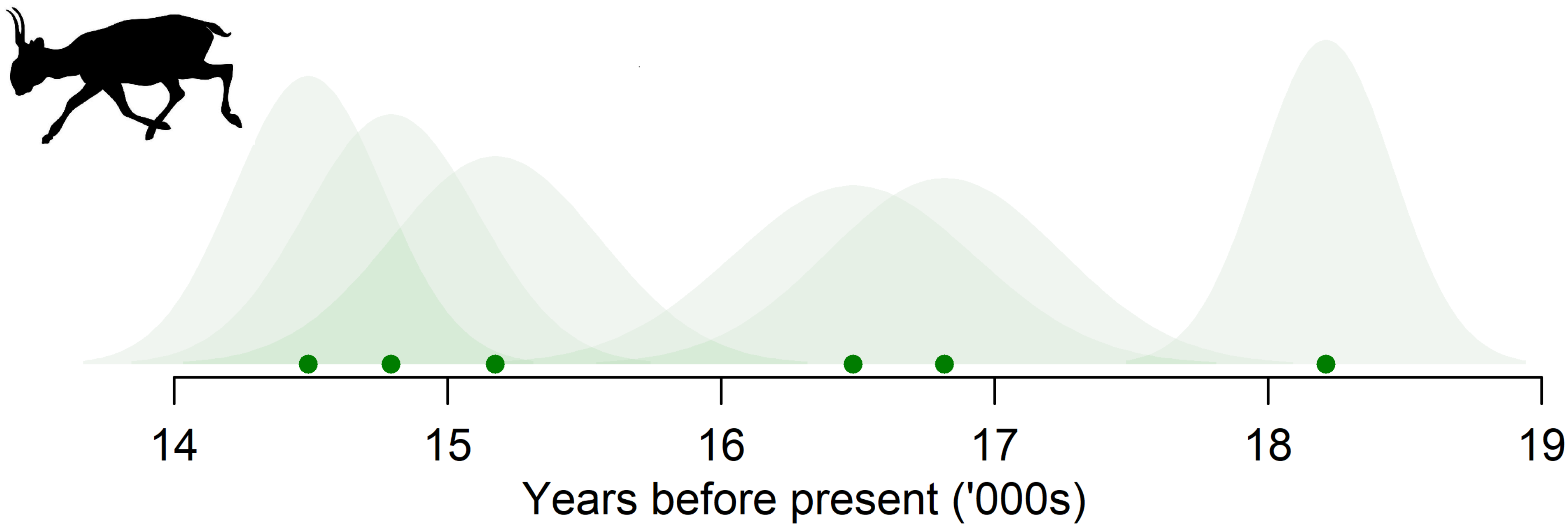
Problem: doesn't work for small σ (or small n)

… and simulations show asymptotic methods are no good for small σ and n.

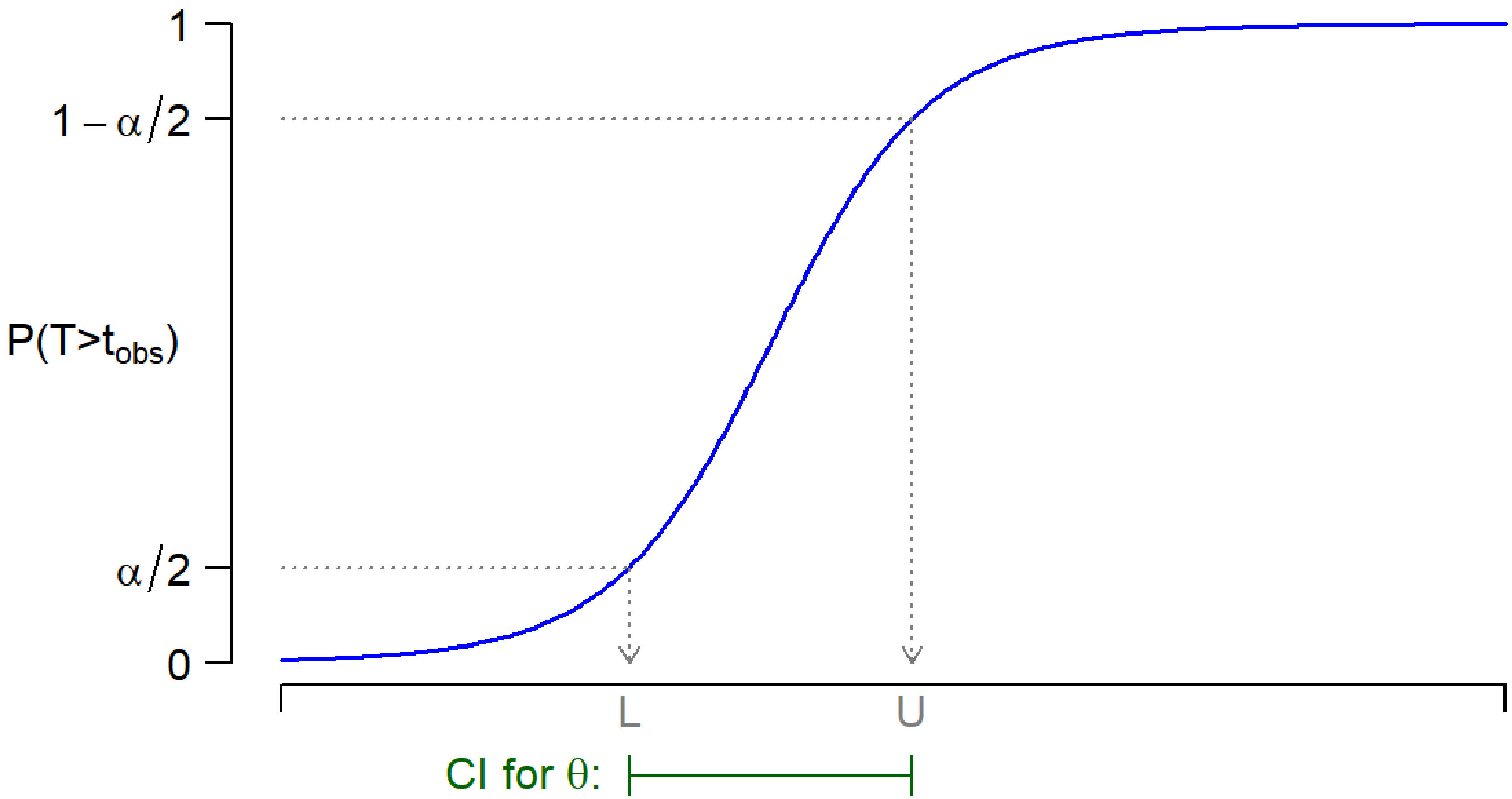


Problem: doesn't work for small σ (or small n)

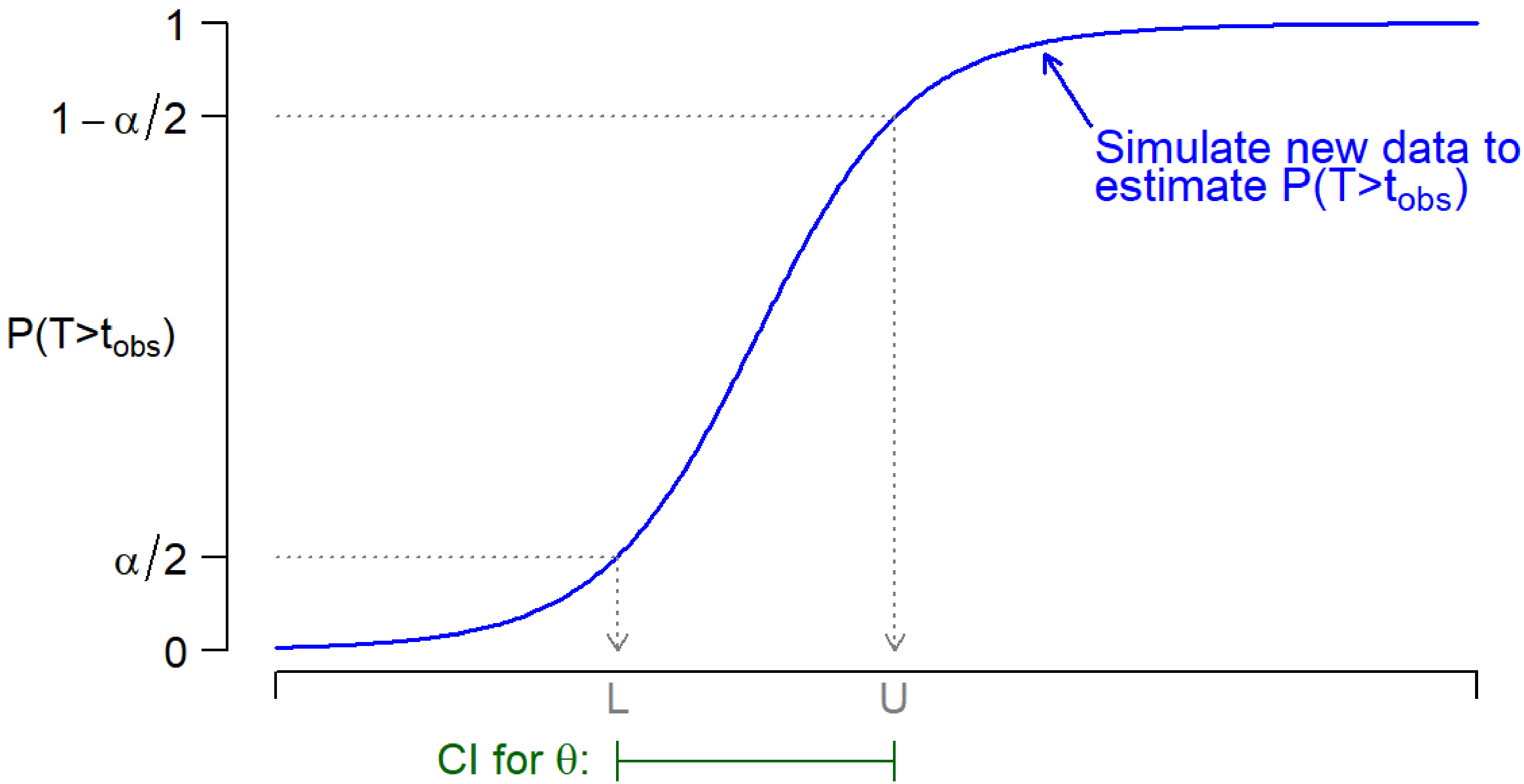
We usually have small σ and n!



Solution:



Solution: simulated (or “bootstrap”) inversion

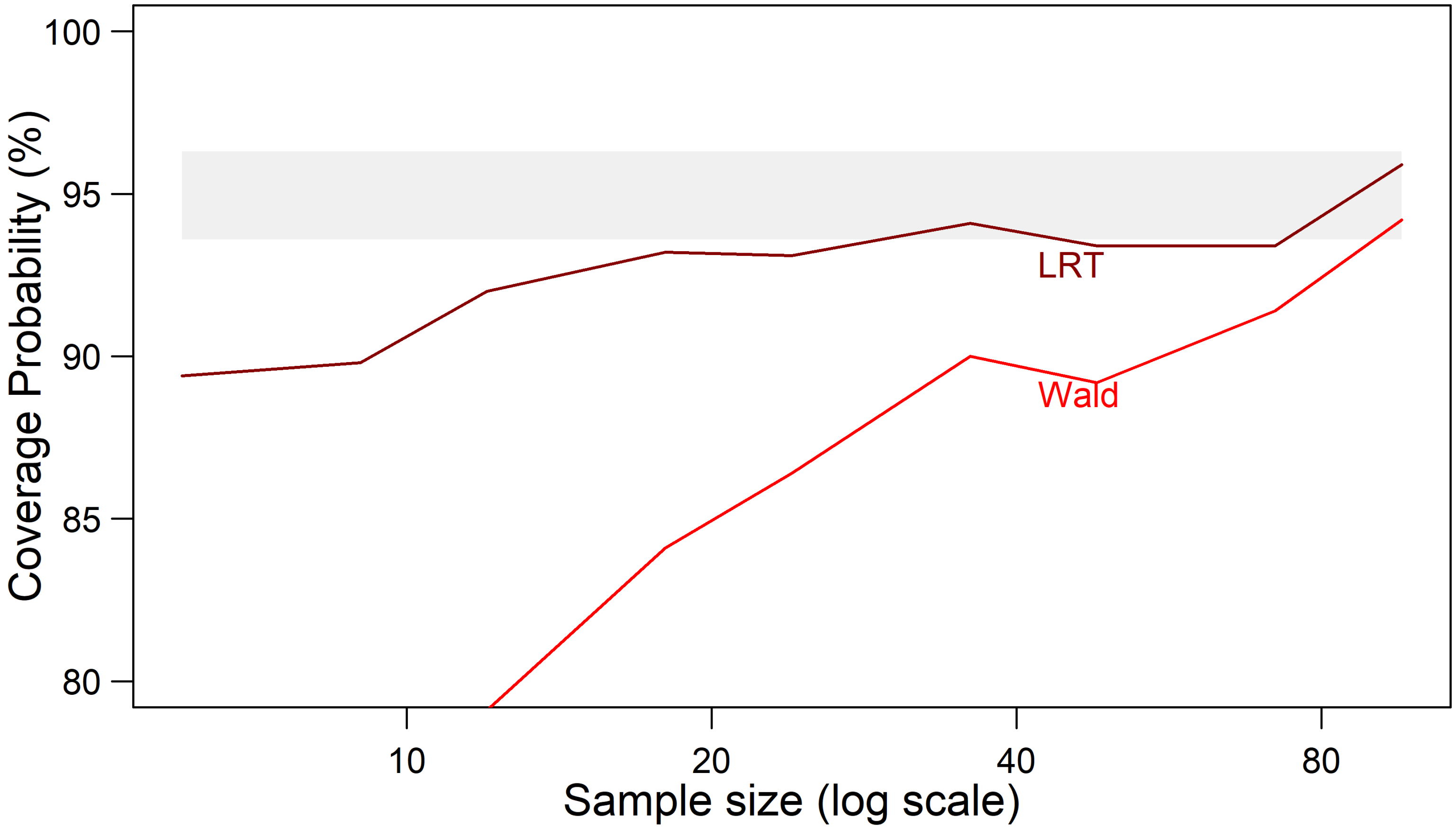


Simulated inversion

We tried two methods:

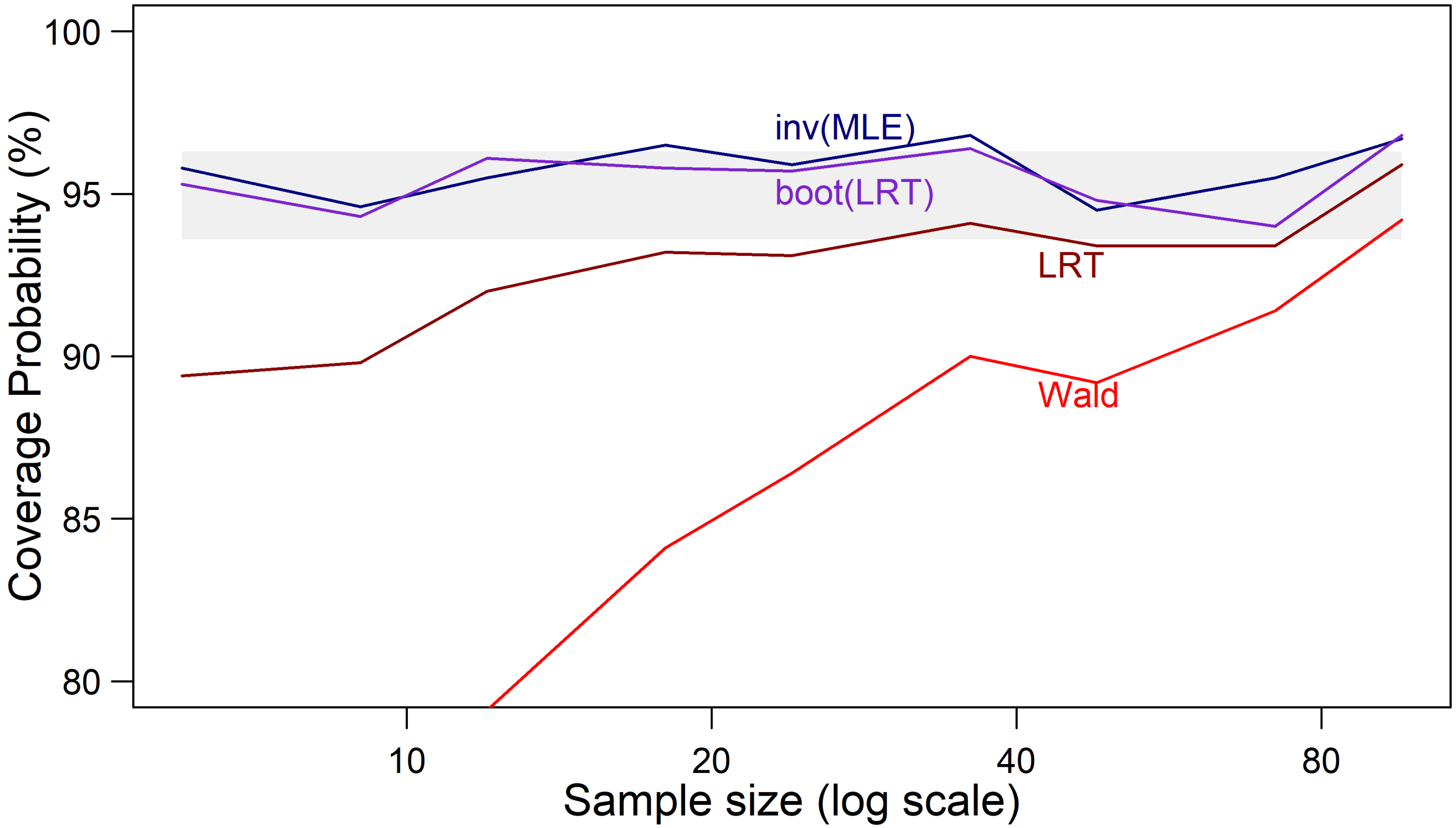
- `inv(MLE)` – use $T = \hat{\theta}$, get quantiles using quantile regression, simulate with θ set to current estimate. Similar to Fisher, Schweiger, and Rosset (2020).
- `boot(LRT)` – use $T = \mathrm{sign}(\hat{\theta} - \theta)\sqrt{-2\log\Lambda(\theta)}$, simulate at asymptotic estimates for CI limits and get sample quantiles (*parametric bootstrap*)

Results:

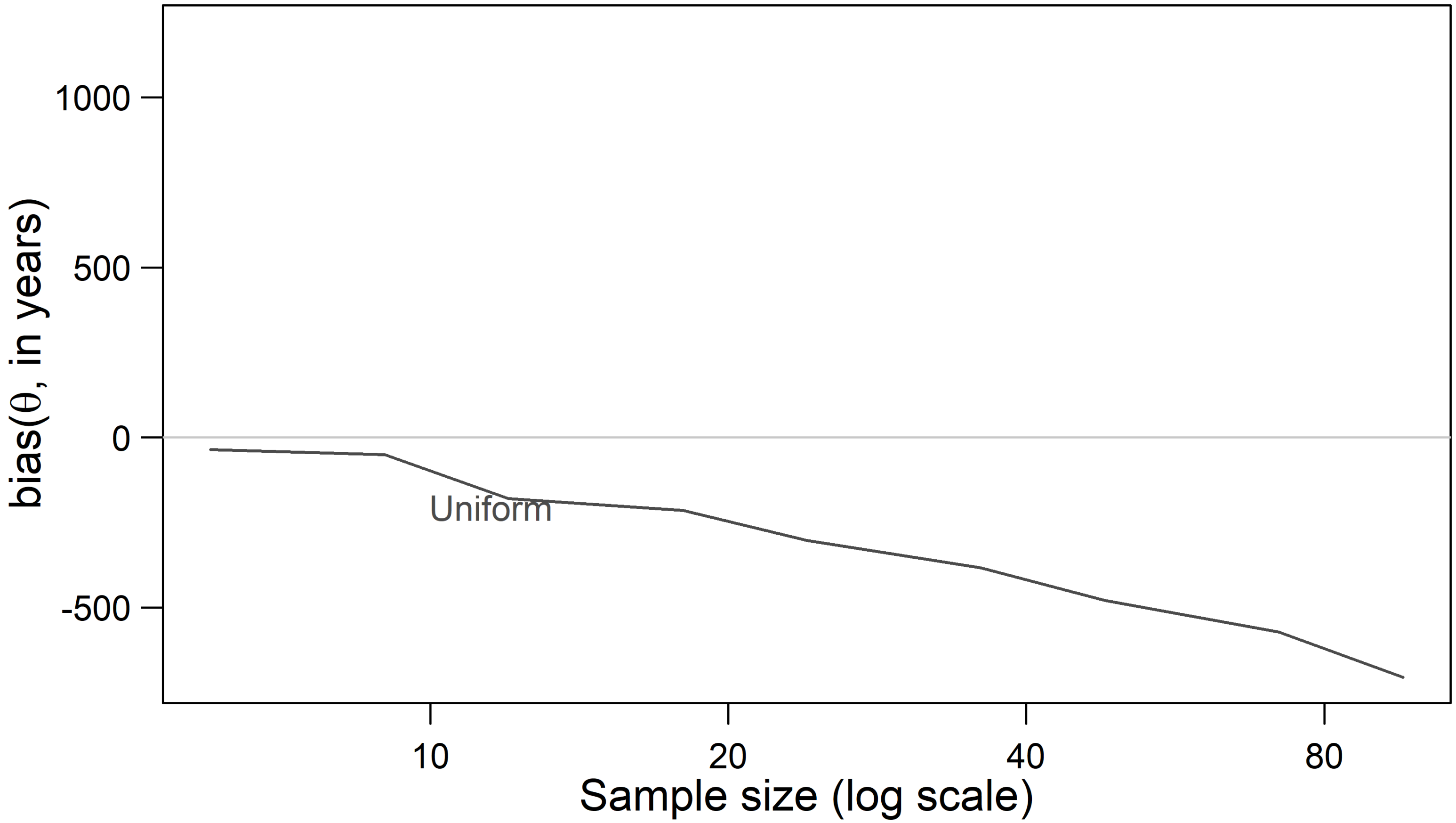


Coverage Probability (%)
100
95
90
85
80
LRT
Wald
10
20
40
80
Sample size (log scale)

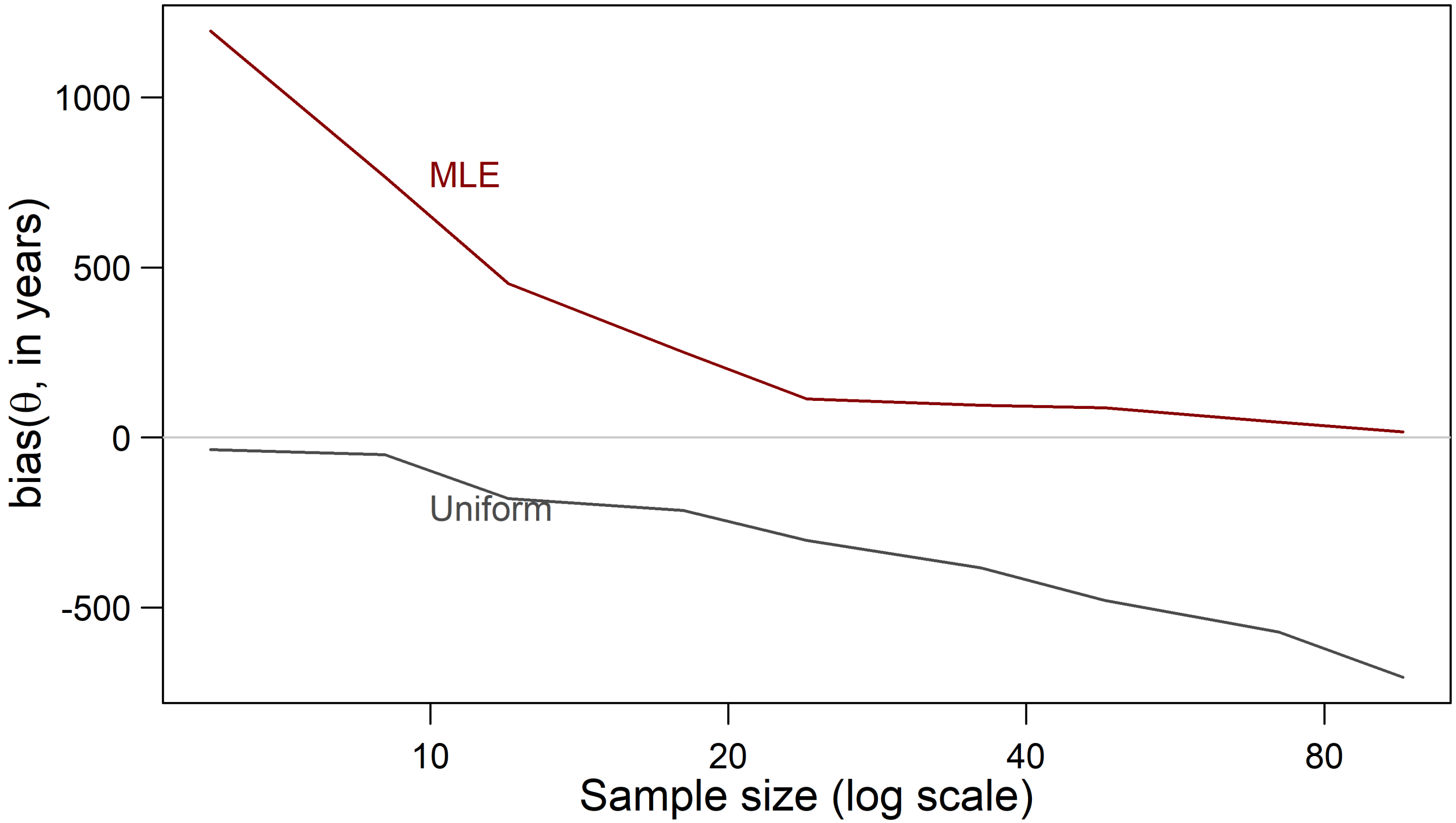
Results: good coverage for small σ or n



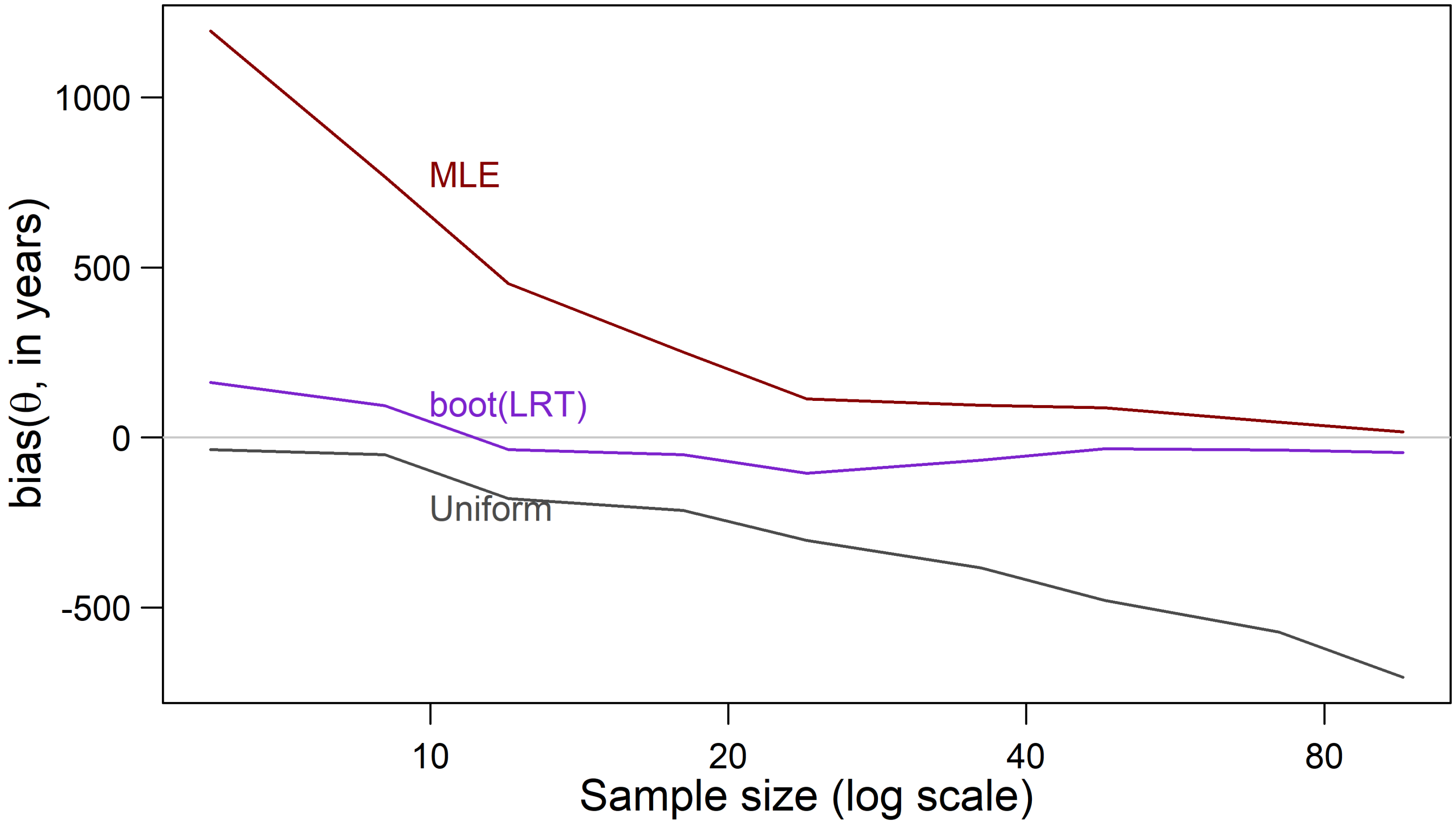
Results: bias



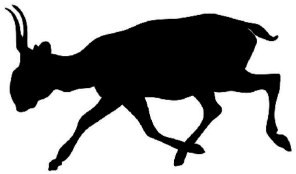
Results: bias correction needed…

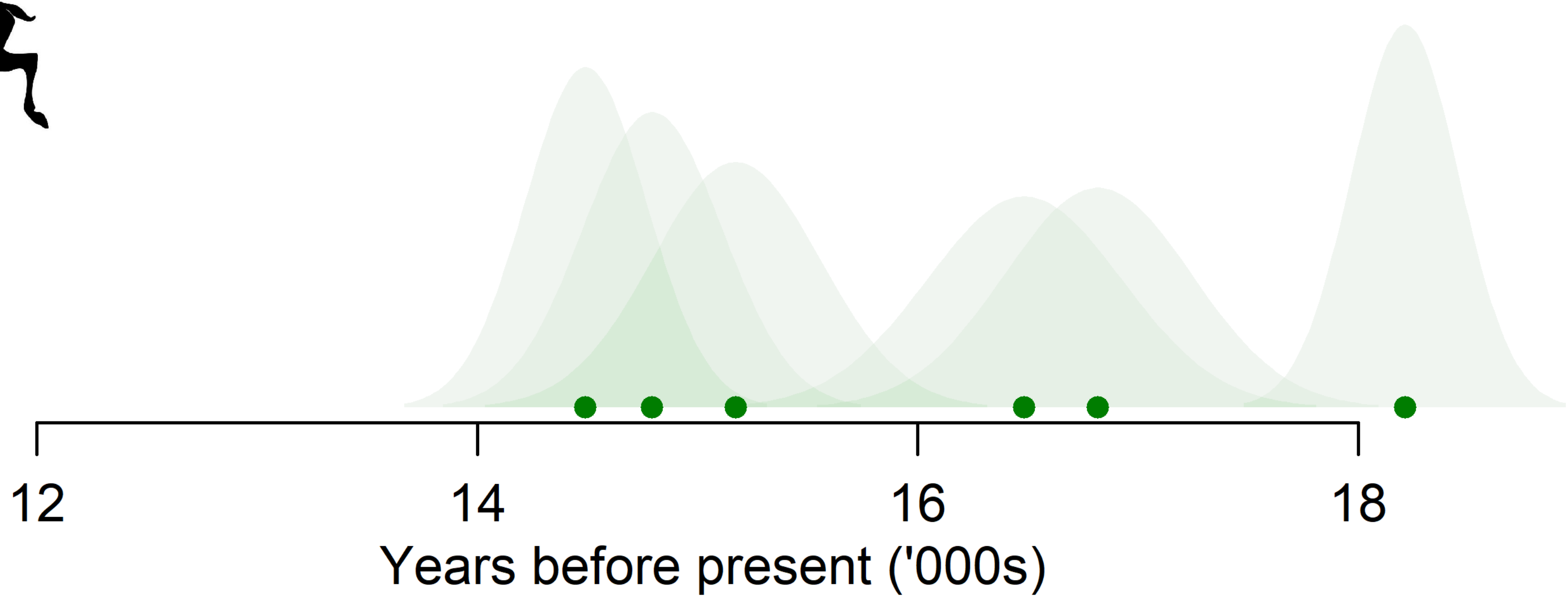


Results: bias correction needed…

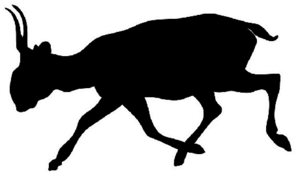


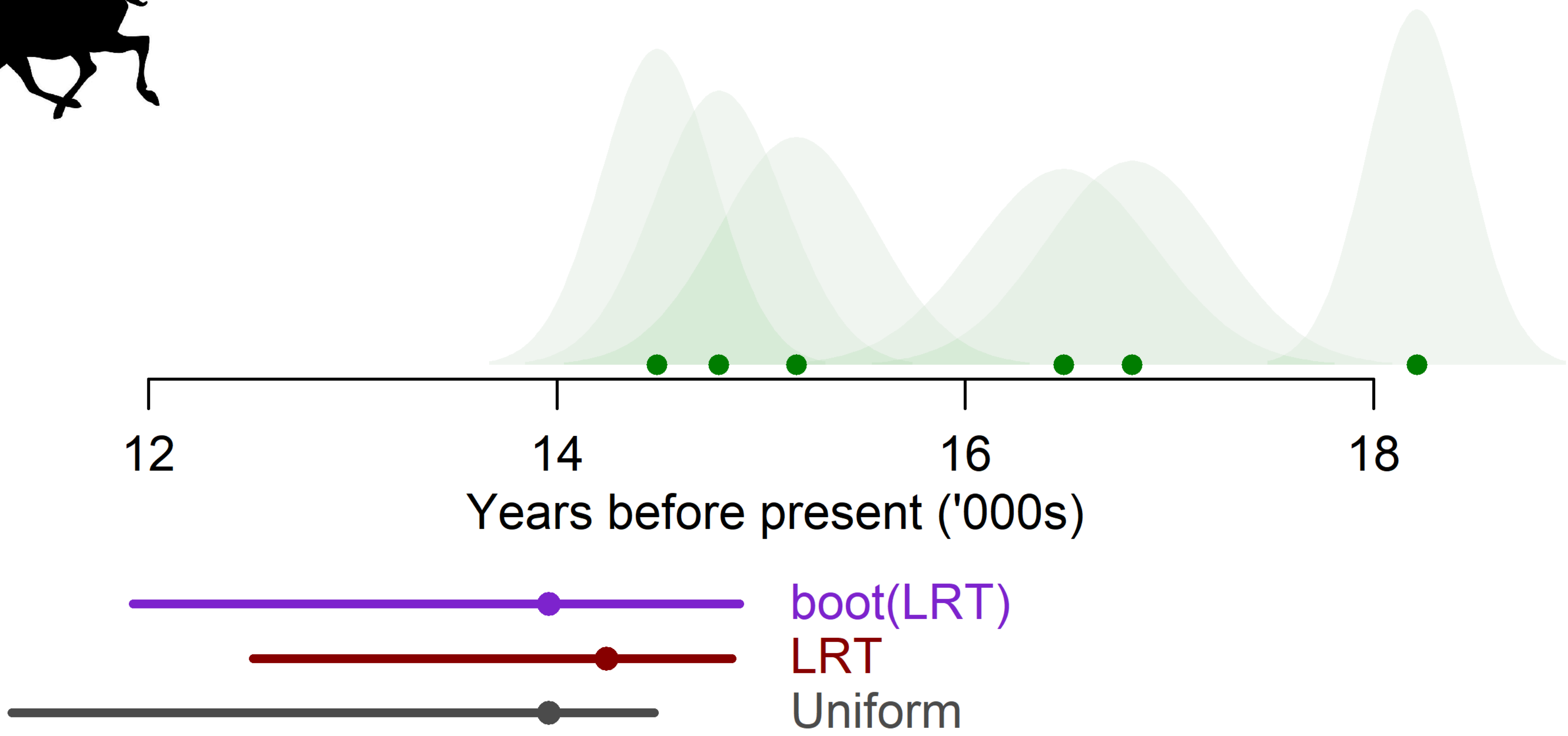
Results: Saiga Antelope extinction estimates





Results: Saiga Antelope extinction estimates





Conclusions

When estimating extinction time from the fossil record:

- You need to account for sampling and measurement error
- Common methods ignore measurement error $\rightarrow$ negative bias
- Classical asymptotics not a good idea $\rightarrow$ bootstrap from asymptotic estimates
- Future work to consider relaxing uniform assumption…

Acknowledgements

- Chris Turney and Alan Cooper for data
- UNSW Eco-Stats group for feedback
- UNSW School of Mathematics and Statistics

This research was undertaken on land traditionally owned by the Bedegal and Wallumedegal people, who I would like to acknowledge and pay my respect to, as well as to Aboriginal Elders past, present and emerging.

References

Bradshaw, C. J. A., A. Cooper, C. S. M. Turney, and B. W. Brook. 2012. "Robust Estimates of Extinction Time in the Geological Record." *Quaternary Science Reviews* 33 (February): 14–19.

Fisher, Eyal, Regev Schweiger, and Saharon Rosset. 2020. "Efficient Construction of Test Inversion Confidence Intervals Using Quantile Regression." *Journal of Computational and Graphical Statistics* 29 (1): 140–48.

Solow, Andrew R., David L. Roberts, and Karen M. Robbirt. 2006. "On the Pleistocene Extinctions of Alaskan Mammoths and Horses." *Proceedings of the National Academy of Sciences of the United States of America* 103 (19): 7351–53.

Strauss, David, and Peter M. Sadler. 1989. "Classical Confidence Intervals and Bayesian Probability Estimates for Ends of Local Taxon Ranges." *Mathematical Geology* 21 (4): 411–27.