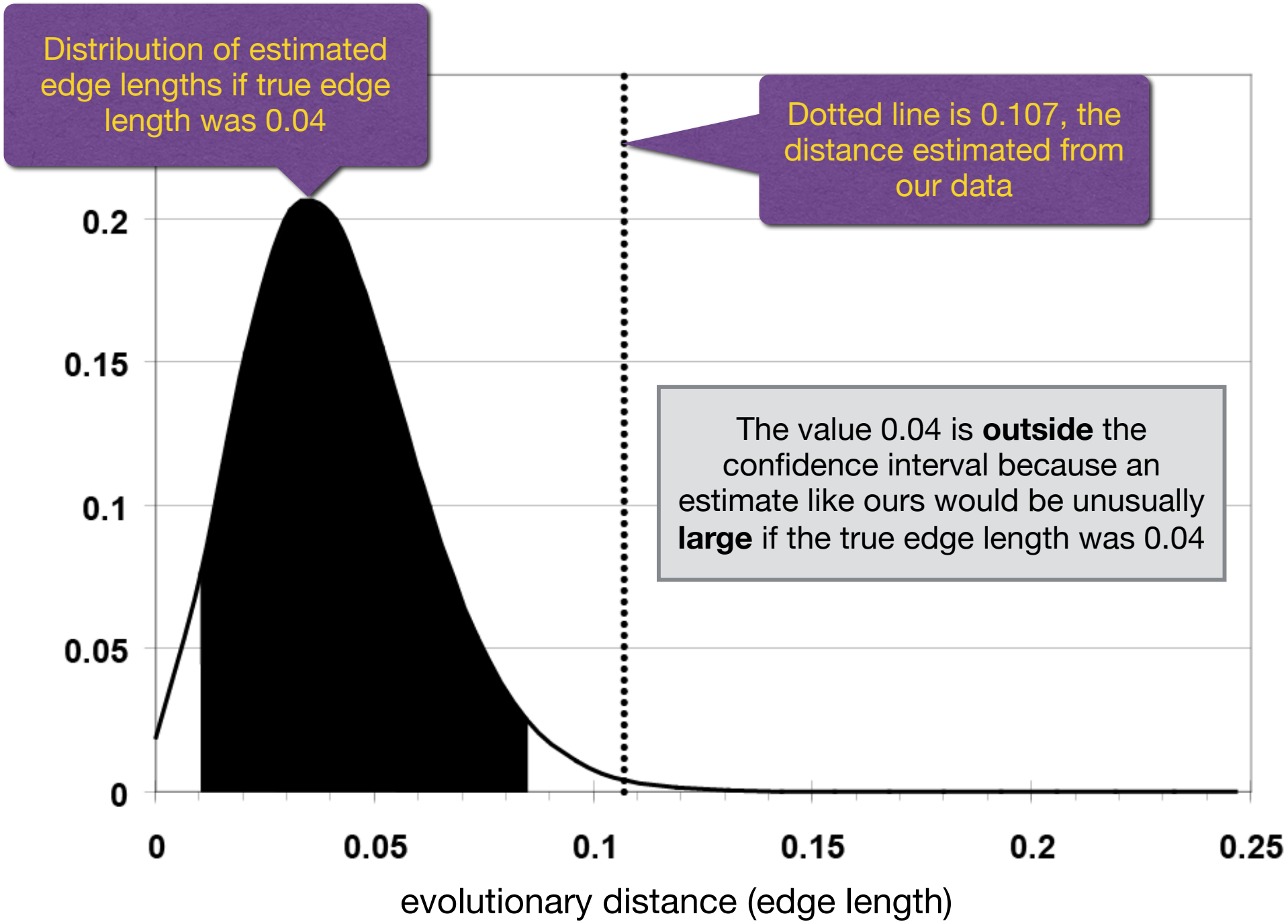
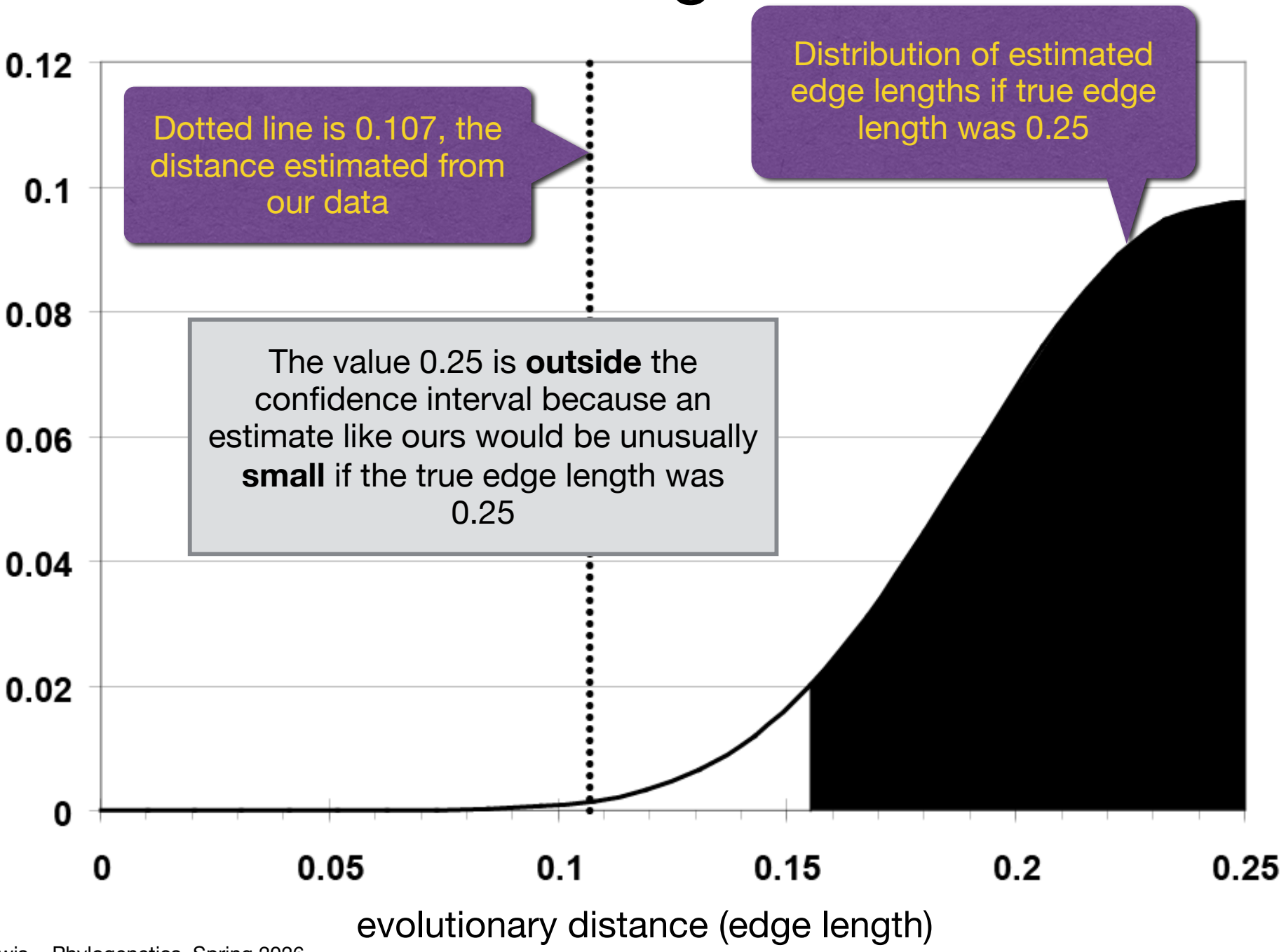
Confidence intervals
vs.
Credible intervals

Confidence interval example

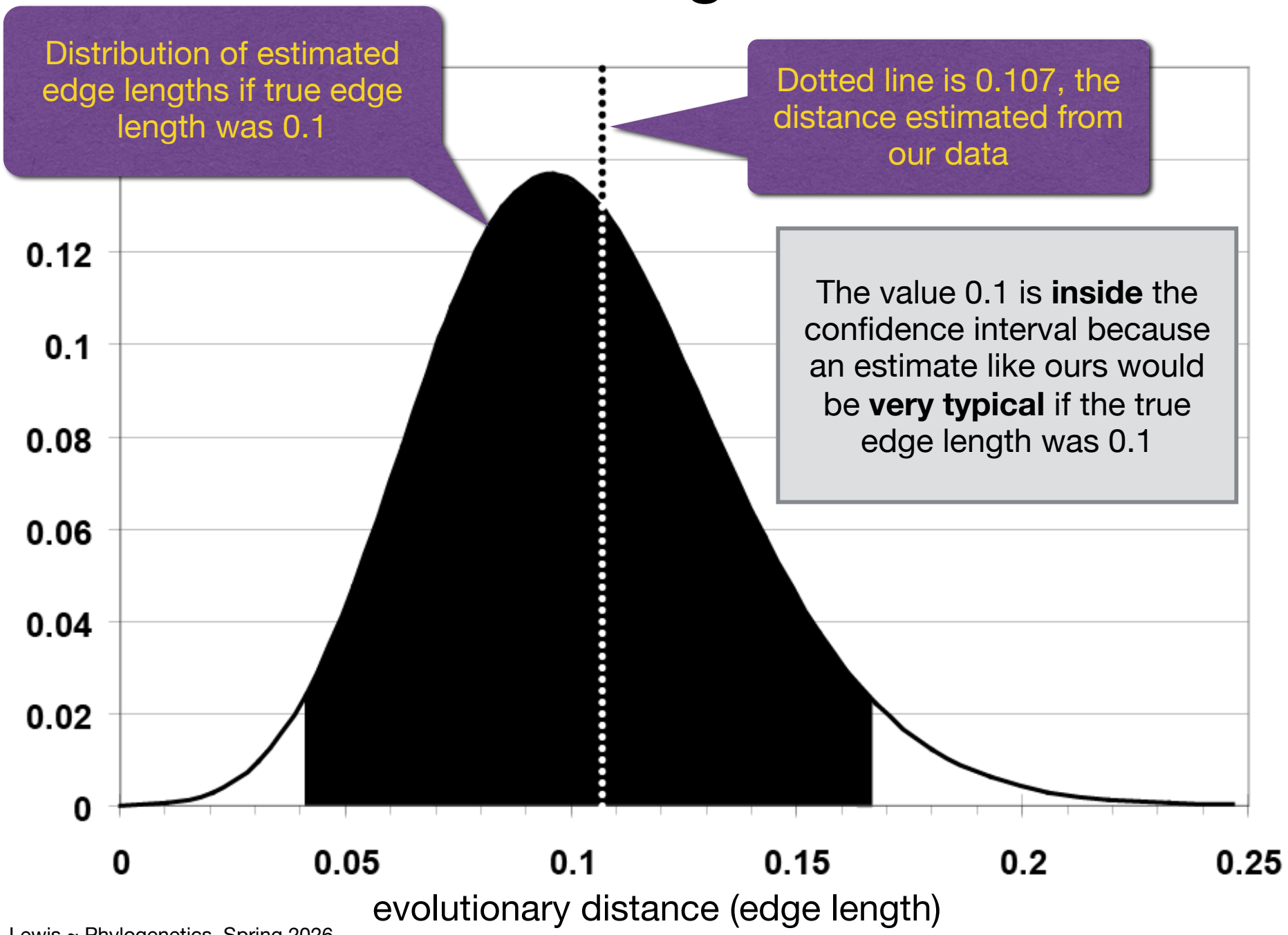
- Interested in the evolutionary distance between two sequences
- 10 differences out of 100 sites leads to JC estimate of 0.107 expected substitutions per site
- How much confidence can we have in this value? A 95% confidence interval is constructed (0.06, 0.19), but what does this interval mean?

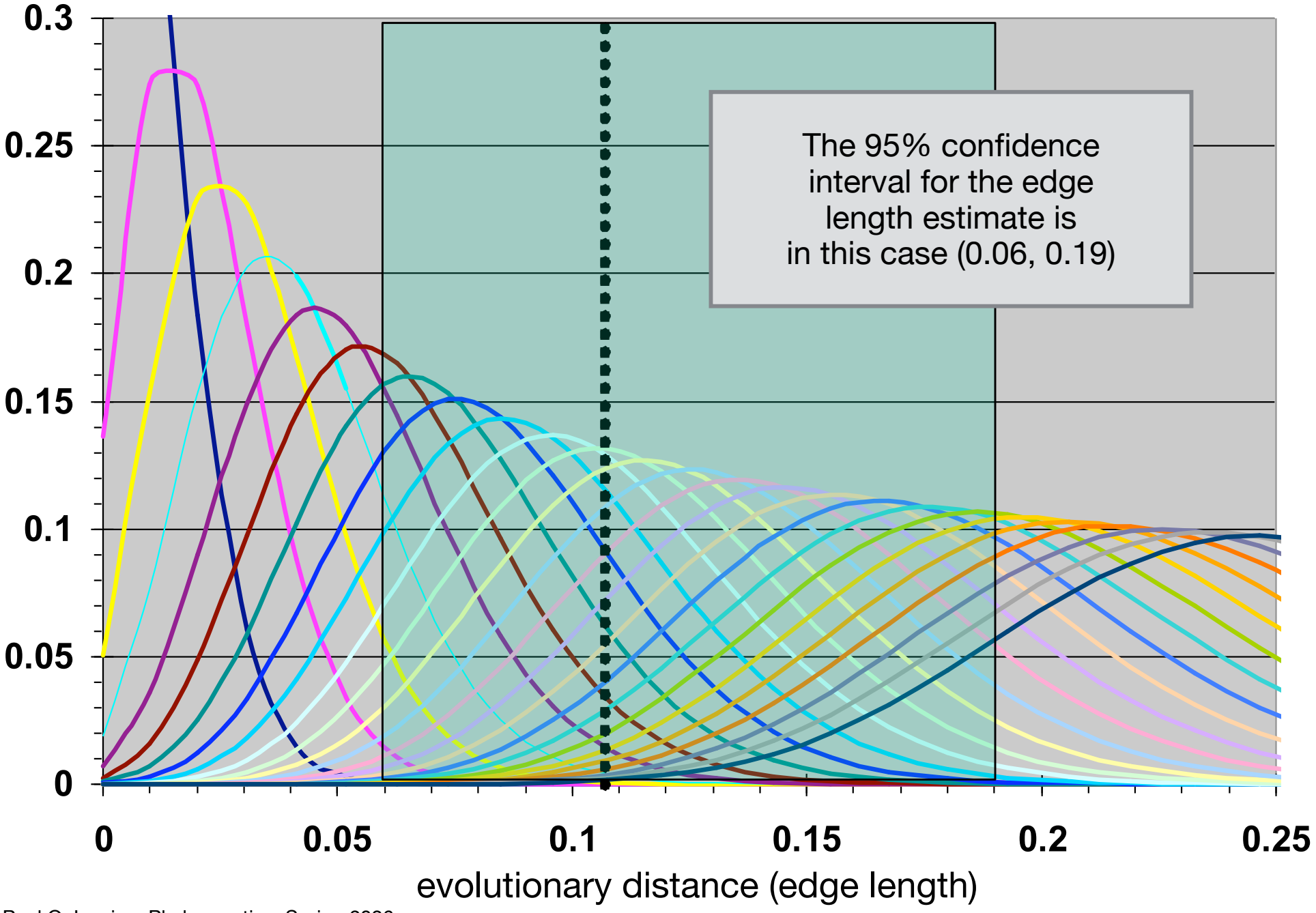

Distribution of estimated edge lengths if true edge length was 0.04
Dotted line is 0.107, the distance estimated from our data
The value 0.04 is **outside** the confidence interval because an estimate like ours would be unusually **large** if the true edge length was 0.04
0.2
0.15
0.1
0.05
0
0
0.05
0.1
0.15
0.2
0.25
evolutionary distance (edge length)

Too high


Distribution of estimated edge lengths if true edge length was 0.25
Dotted line is 0.107, the distance estimated from our data
The value 0.25 is **outside** the confidence interval because an estimate like ours would be unusually **small** if the true edge length was 0.25
0.12
0.1
0.08
0.06
0.04
0.02
0
0
0.05
0.1
0.15
0.2
0.25
evolutionary distance (edge length)

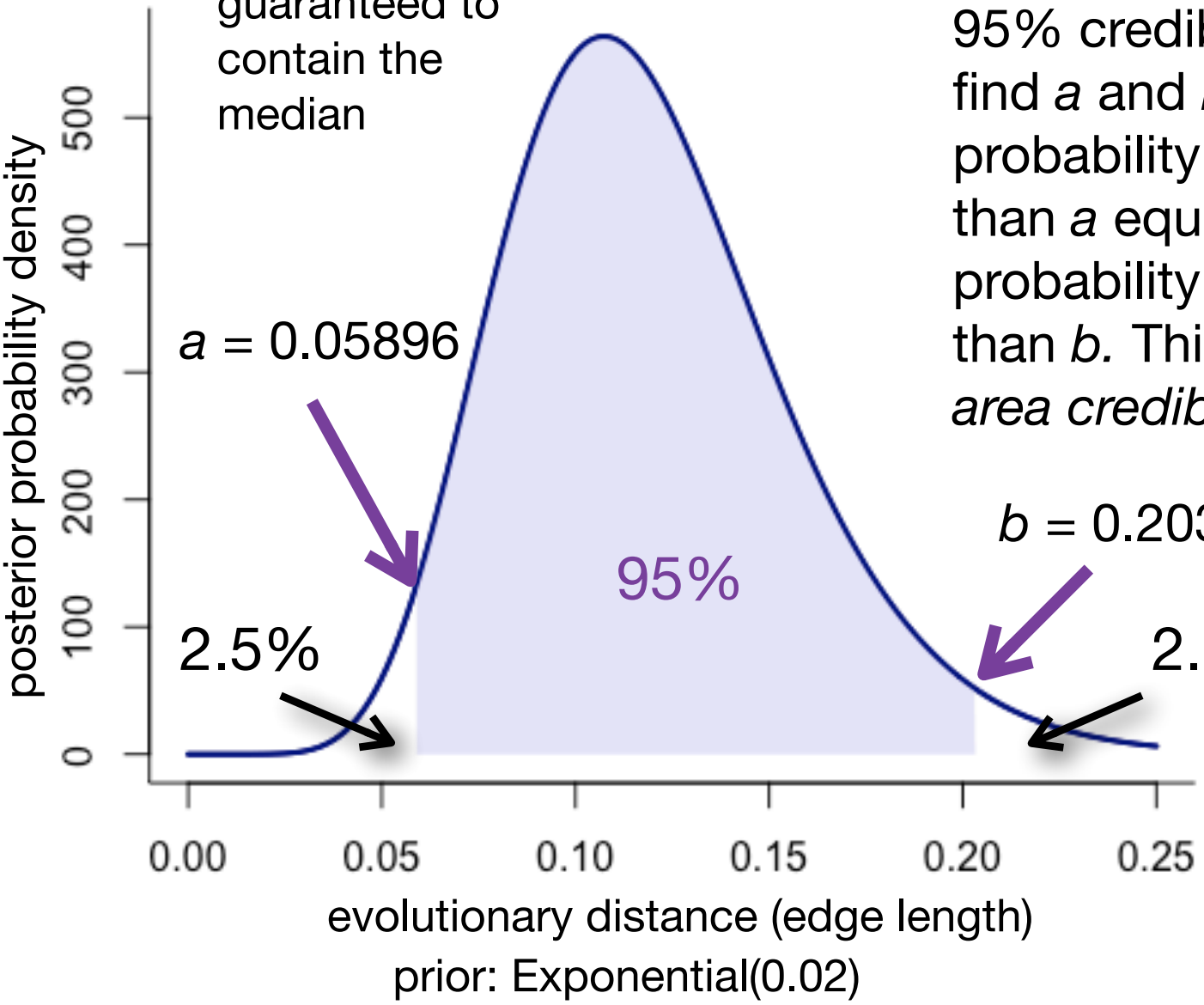
Just right


Distribution of estimated edge lengths if true edge length was 0.1
Dotted line is 0.107, the distance estimated from our data
The value 0.1 is **inside** the confidence interval because an estimate like ours would be **very typical** if the true edge length was 0.1
0.12
0.1
0.08
0.06
0.04
0.02
0
0
0.05
0.1
0.15
0.2
0.25
evolutionary distance (edge length)


0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.05
0.1
0.15
0.2
0.25
evolutionary distance (edge length)
The 95% confidence interval for the edge length estimate is in this case (0.06, 0.19)

Credible Intervals

Equal tail area intervals are guaranteed to contain the median

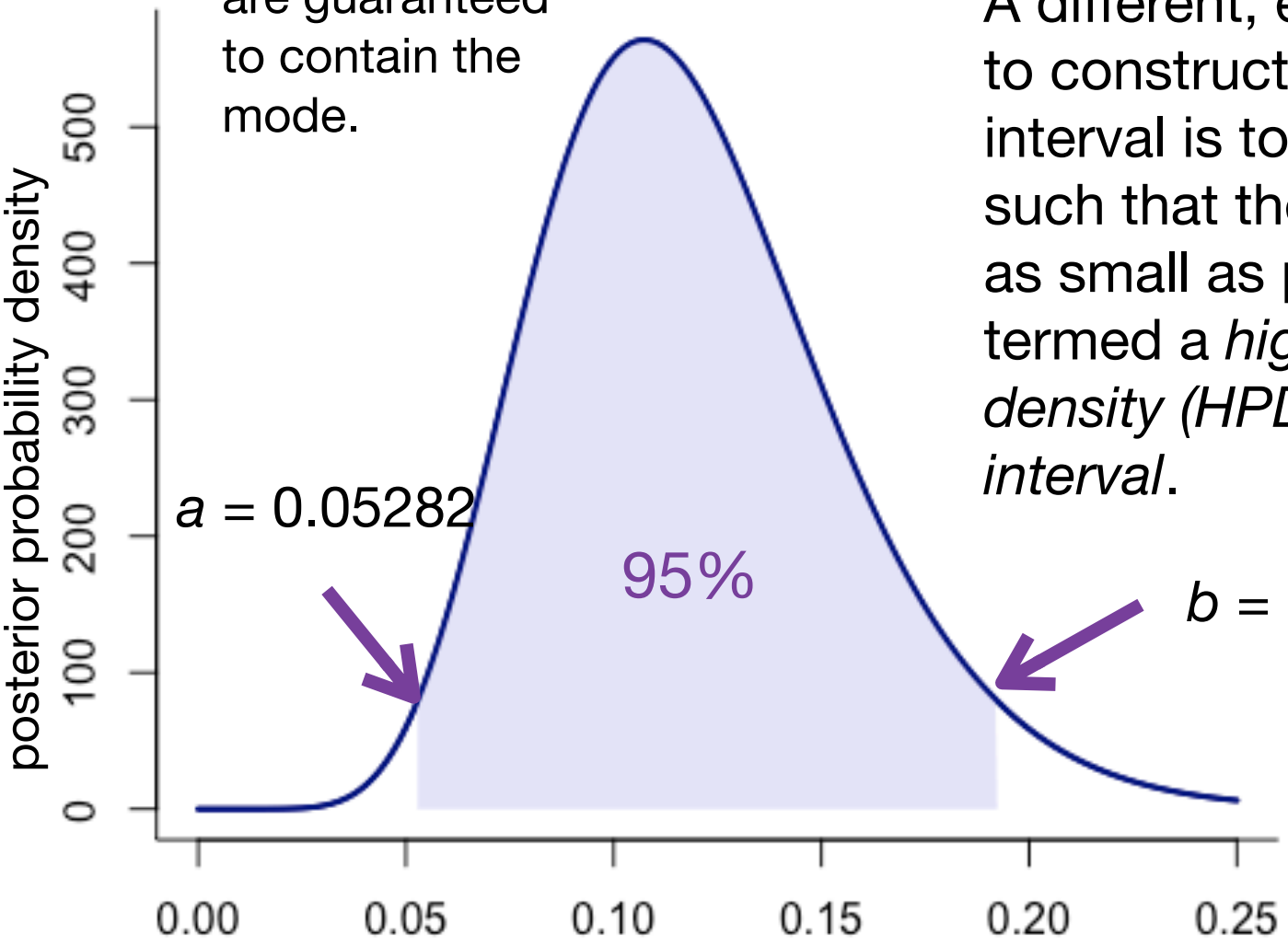


One way to construct a 95% credible interval is to find *a* and *b* such that the probability of being less than *a* equals the probability of being greater than *b.* This is an *equal tail area credible interval*.

$b - a = 0.14423$

HPD credible intervals

HPD intervals are guaranteed to contain the mode.



prior: Exponential(0.02)

A different, equally-valid way to construct a 95% credible interval is to find *a* and *b* such that the difference *b*-*a* is as small as possible. This is termed a *highest probability density (HPD) credible interval*.

b - *a* = 0.13951

(cf. 0.14423 for equal-tail)