



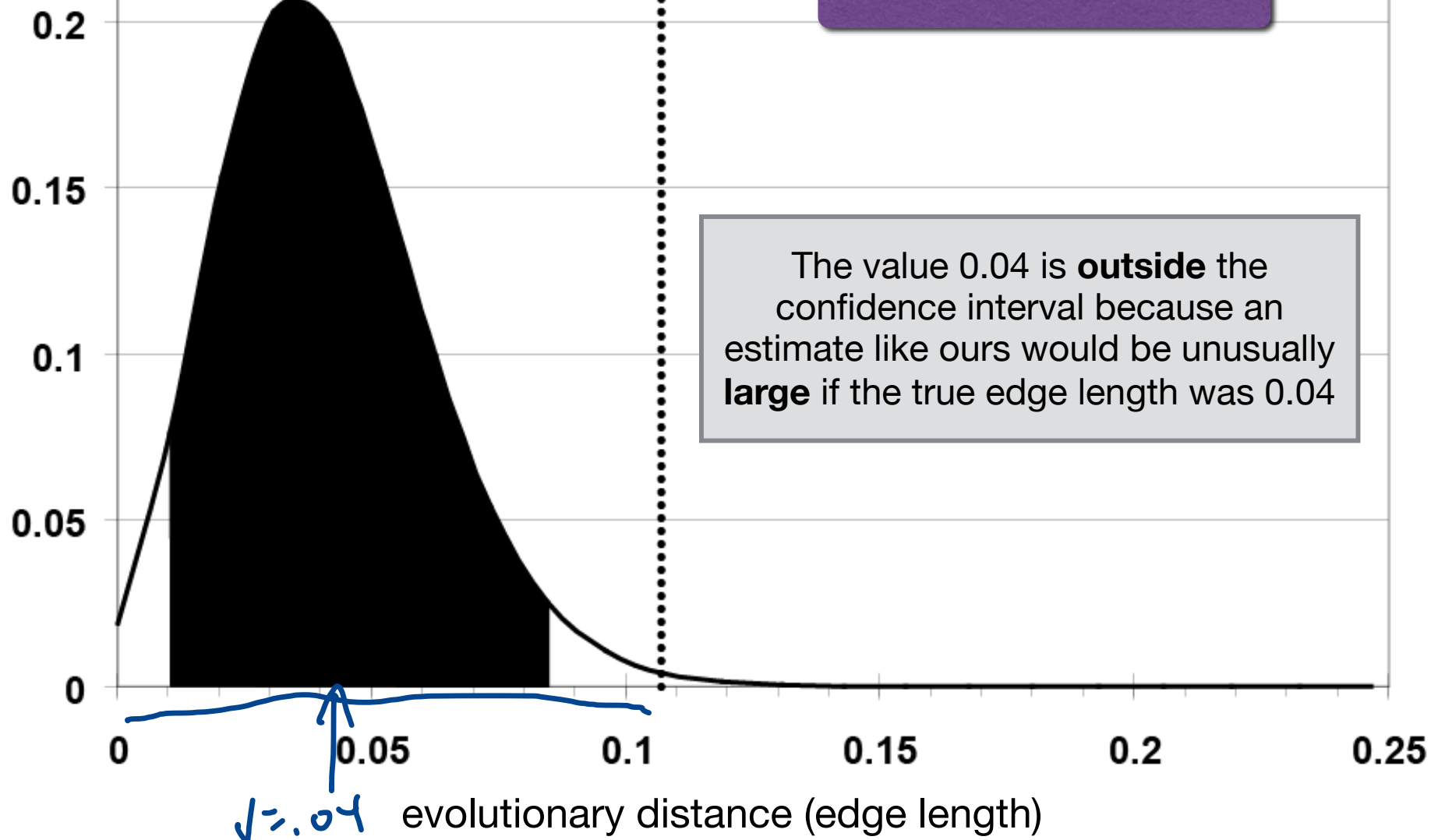
Confidence interval example

- 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?

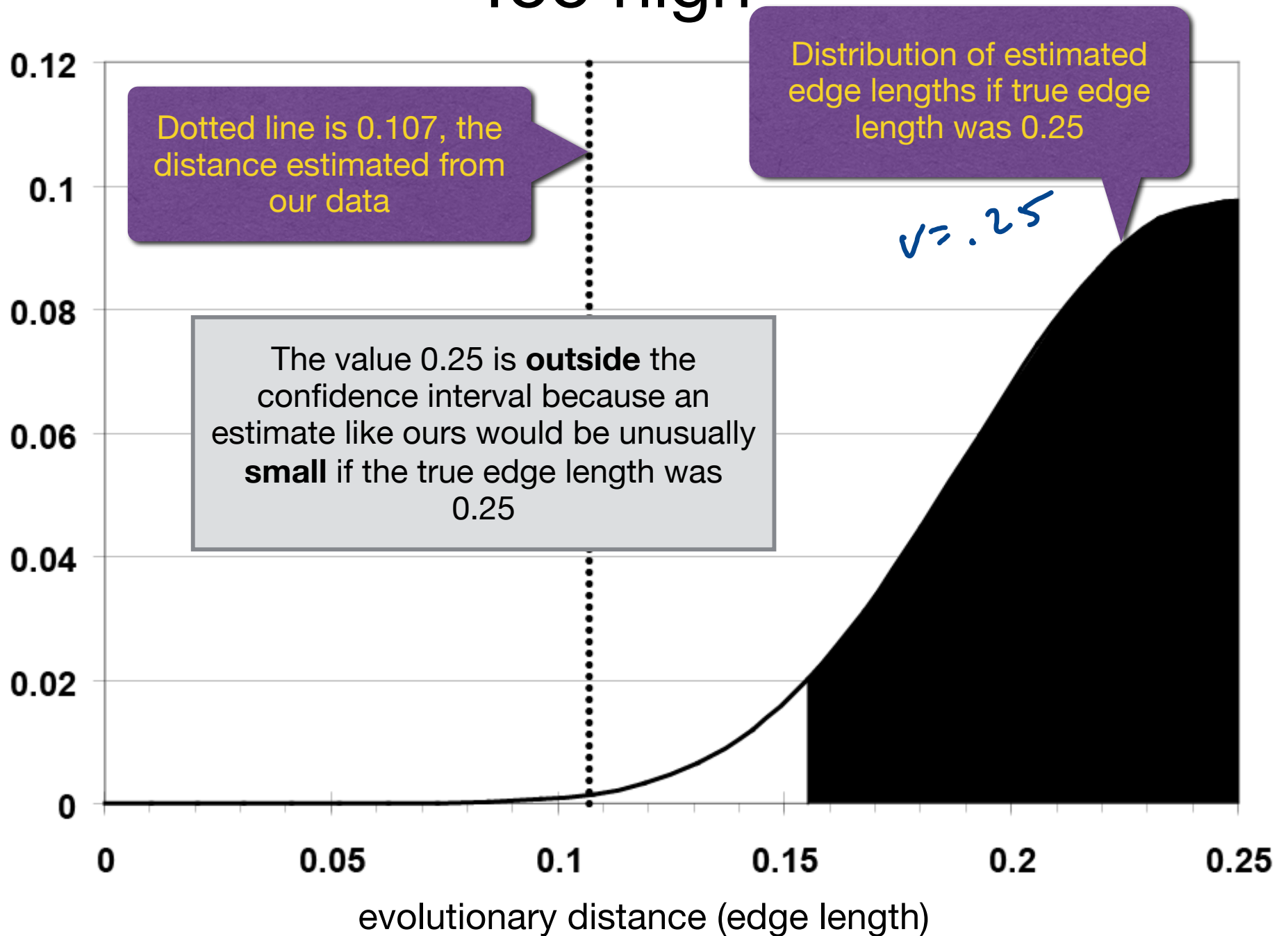
Too low

Distribution of estimated edge lengths if true edge length was 0.04

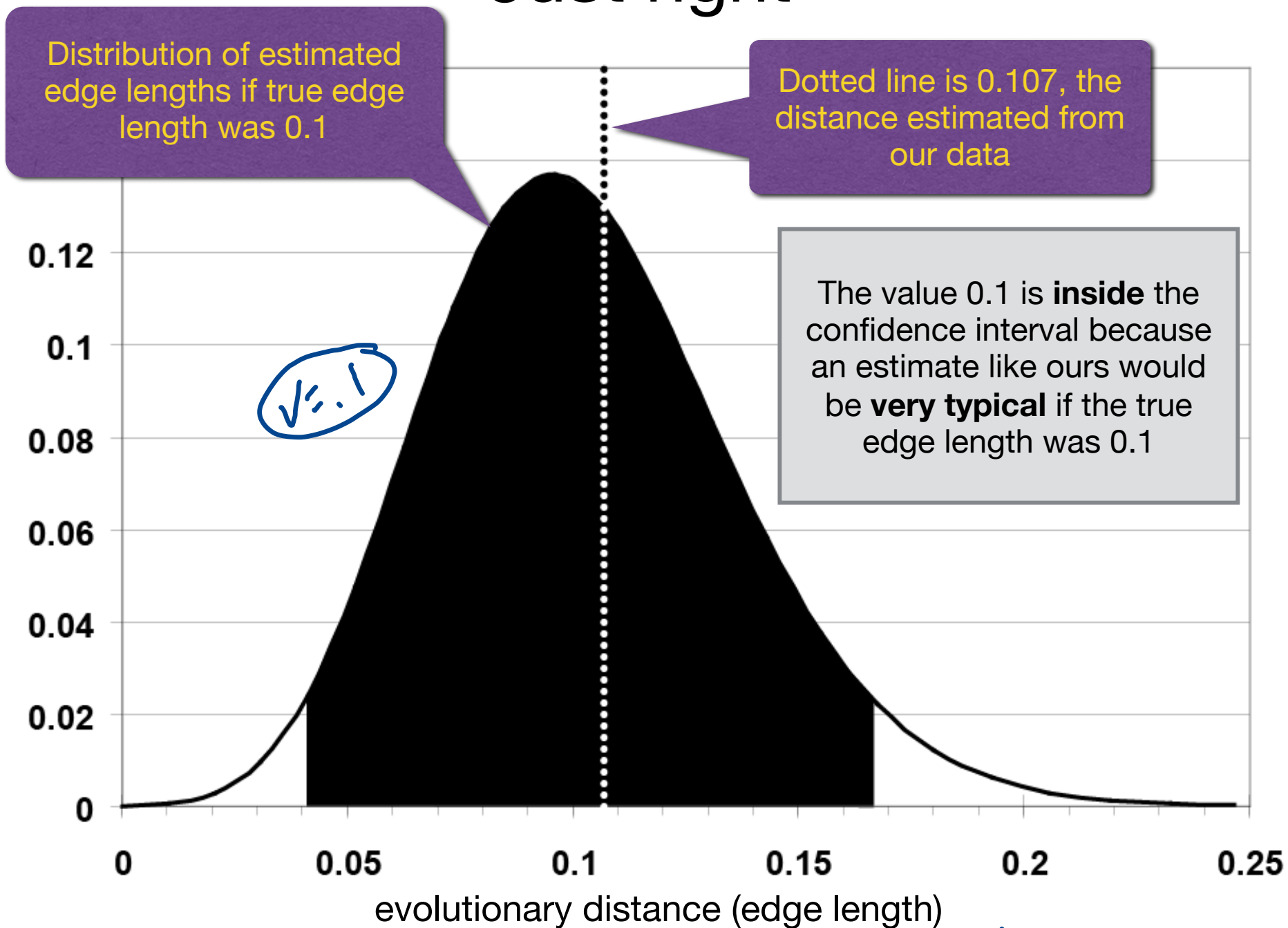
Dotted line is 0.107, the distance estimated from our data



Too high

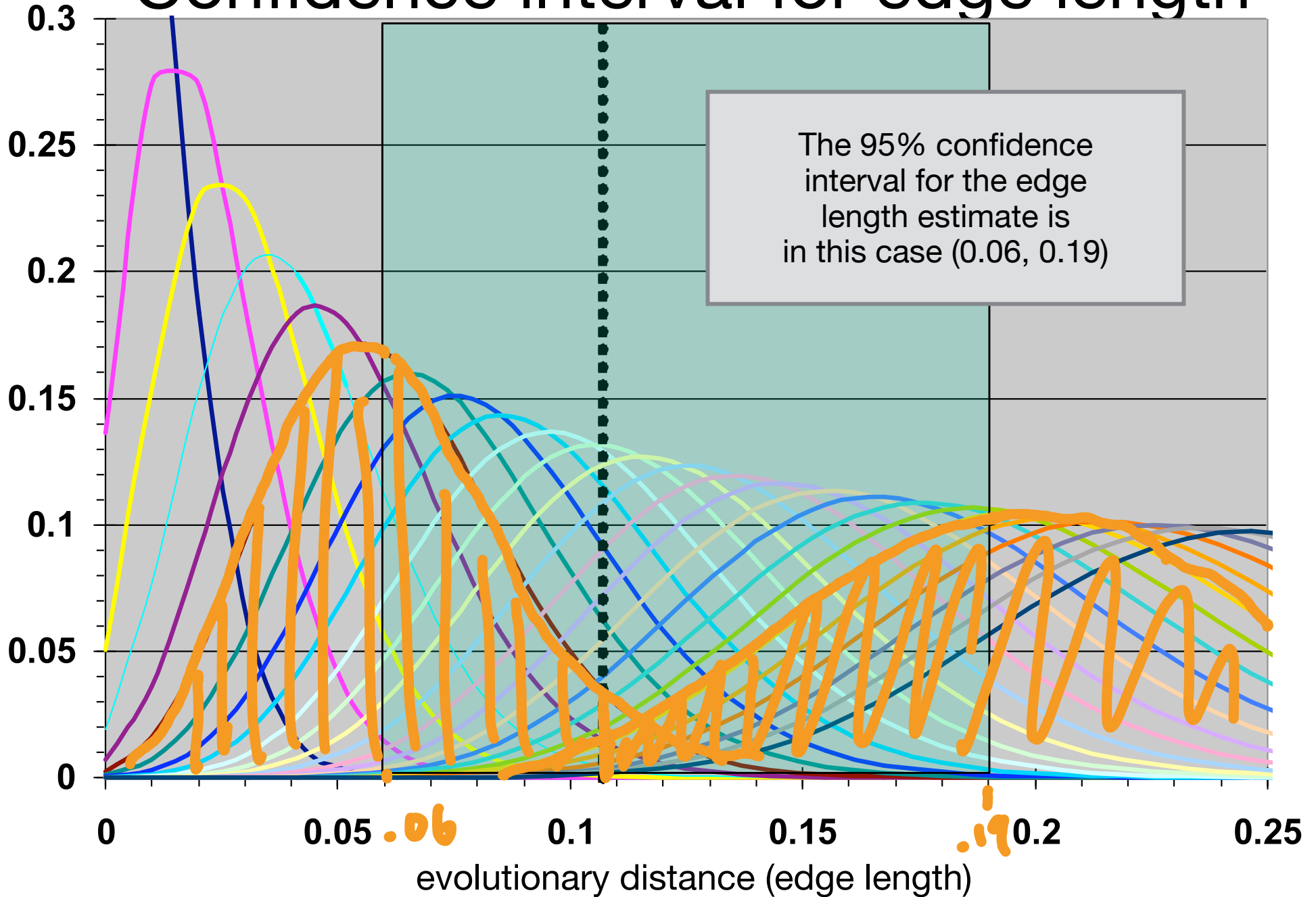


Just right



$p(D|v) \leftarrow \text{likelihood}$

Confidence interval for edge length

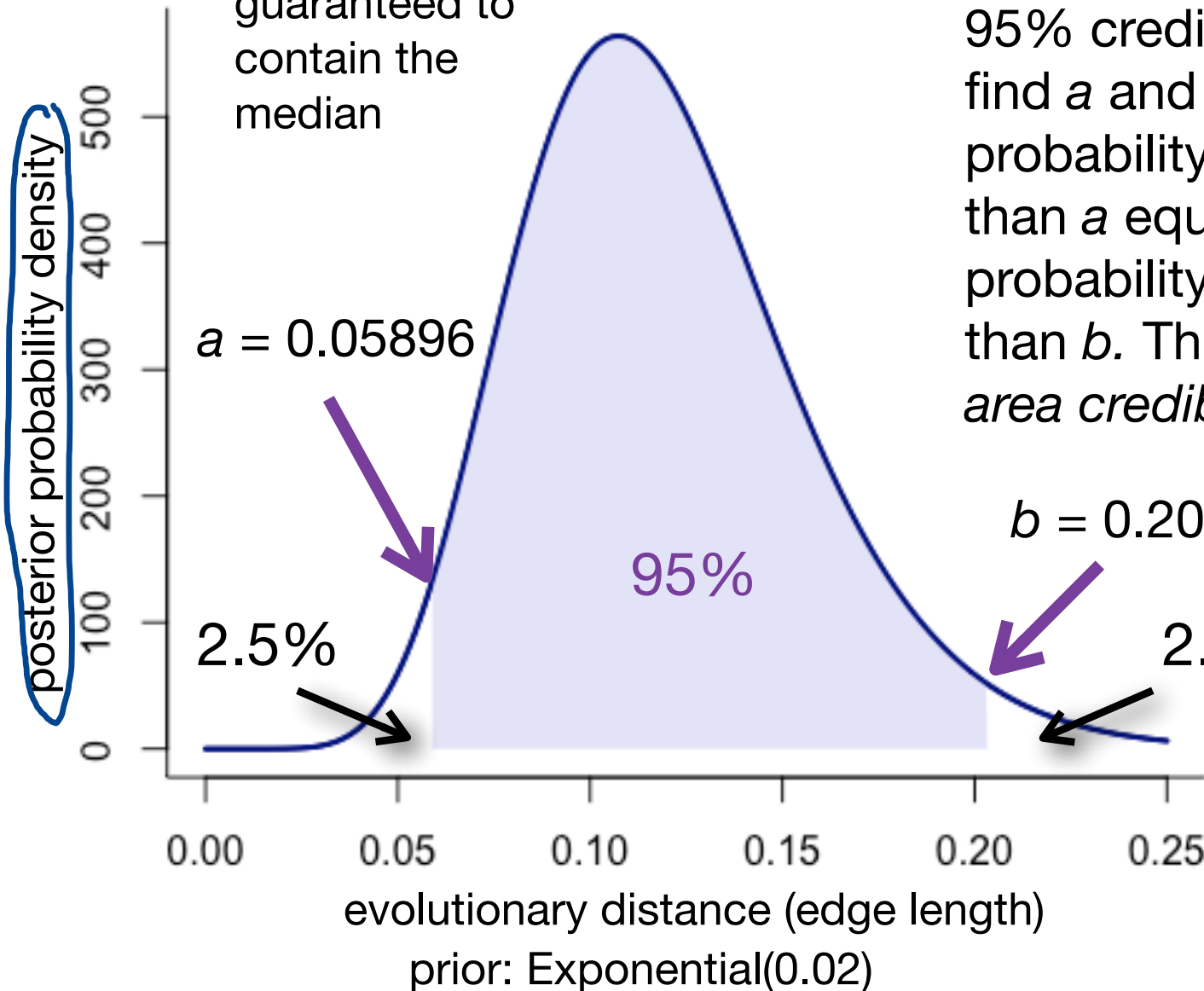


Credible Intervals

$P(V|D)$
↑
variable
 μ fixed

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*.



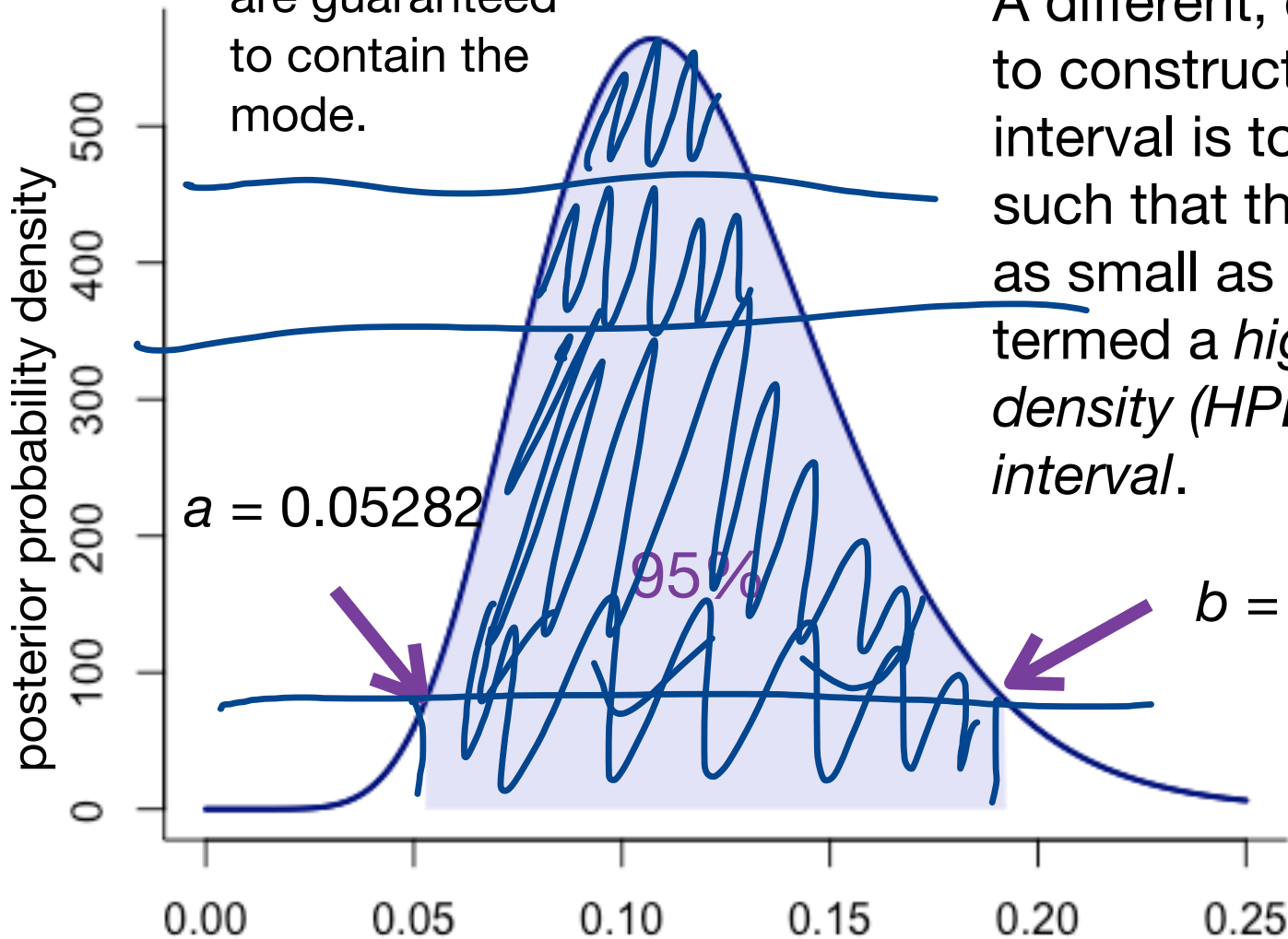
$$b - a = 0.14423$$

Highest Probability Density

HPD credible intervals

HPD intervals are guaranteed to contain the mode.

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*.



evolutionary distance (edge length)

prior: Exponential(0.02)

$$b - a = 0.13951$$

(cf. 0.14423 for equal-tail)

