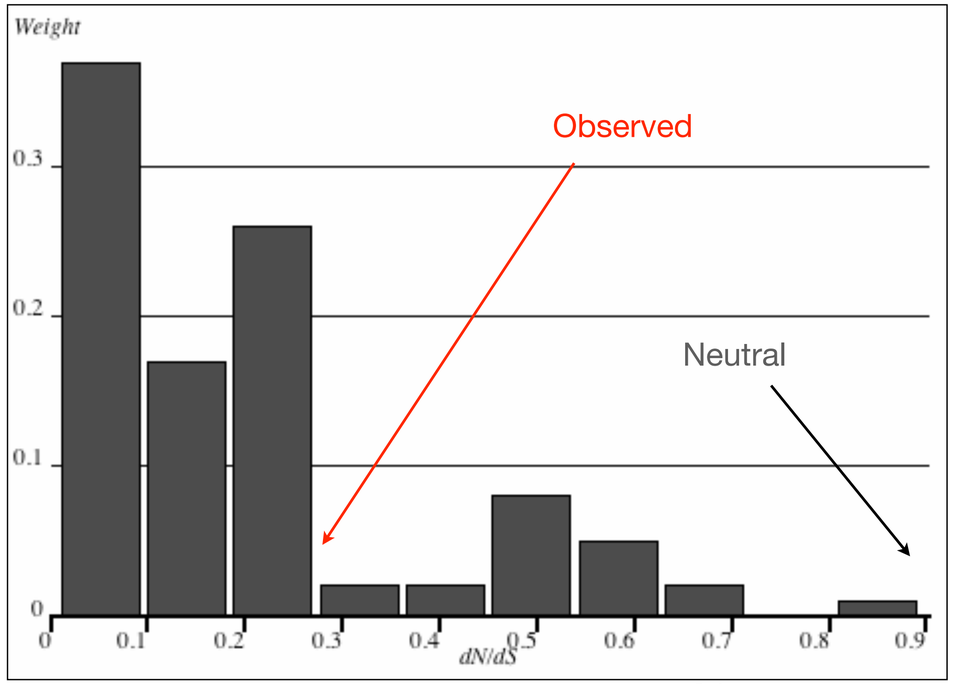
NG86 example

- How reliable is the inference based on only **6** codons?
- Obtain sampling variance via bootstrap (or by limiting approximations)
- In this case, dN/dS is **significantly** less than `1.0 (p ~ 0.01)`

Bootstrapped distribution of dN/dS




```
Count    = 100
Mean     = 0.207385
Median   = 0.166687
Variance = 0.0490168
Std.Dev  = 0.221397
COV      = 1.06757
Sum      = 20.7385
Sq. sum  = 9.15351
Skewness = 0.266313
Kurtosis = 33.381
Min      = 0
2.5%     = 0
97.5%    = 0.741176
Max      = 1
```

NG86 limitations: multiple substitutions

- How many synonymous and how many non-synonymous substitutions does it take to replace **`CCA`** with **`CAG`**?
- **Assume** the shortest path (minimum of 2 substitutions)
  - **`CCA (Proline)`** ⟹ **`CAA (Histidine)`** ⟹ **`CAG (Glutamine)`**
  - **`CCA (Proline)`** ⟹ **`CCG (Proline)`** ⟹ **`CAG (Glutamine)`**
- Average over the two possible paths: `0.5` synonymous and `1.5` non-synonymous substitutions.
- Intuitively, paths should **not** be equiprobable, e.g., because it should be more expensive to route evolution through (presumably) suboptimal intermediate amino-acids.