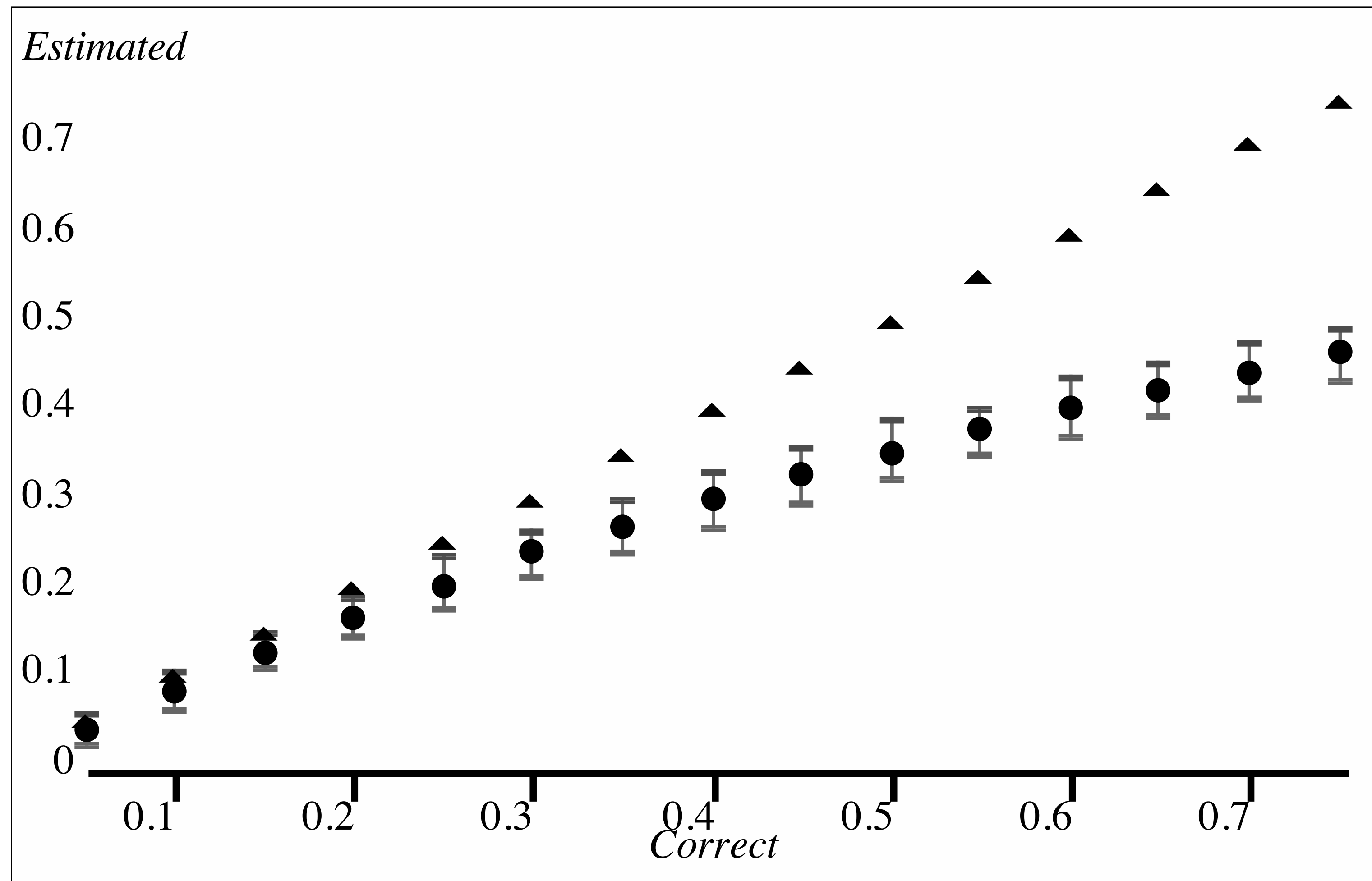


NG86 limitations: underestimation of substitution counts for higher divergence levels

Simulated divergence vs that estimated by p-distance

- Simulated 100 replicates of 1000 nucleotide long sequences for various divergence levels (substitutions/site)
- Even for divergence of 0.25 (1/4 sites have mutation on average), p-distance already underestimates the true level: 0.2125 (0.19–0.241 95% range)
- Underestimation becomes progressively worse for larger divergence levels



Nei with his father



1934-1935

Source: Nei Memoir (2000)