Assignment Nr. 10

(Deadline 26.01.2020)

Task 1

$P = 6 \cdot 2 + 3 \cdot 5 + 5 \cdot 3 + 5 \cdot 3 + 6 \cdot 2 + 7 \cdot 1 + 7 \cdot 1 + (1 \cdot 7 + 5 \cdot 2) = 100$
No site is homogenous $\Rightarrow S = 8$
$\Theta_\pi = \frac{P}{\binom{n}{2}} = \frac{100}{\binom{8}{2}} = \frac{100}{28} \approx 3.57$
$\Theta_W = \frac{S}{\sum_{i=1}^{n-1} \frac{1}{i}} = \frac{8}{\sum_{i=1}^{7} \frac{1}{i}} = \frac{8}{\frac{1089}{420}} = \frac{3360}{1089} \approx 3.085$

Task 2

$a_1 = \sum_{i=1}^{n-1} \frac{1}{i} = 1 + ... + \frac{1}{8} \approx 2.59$
$a_2 = \sum_{i=1}^{n-1} \frac{1}{i^2} = 1 + ... + \frac{1}{8^2} \approx 1.51$
$b_1 = \frac{n+1}{3(n-1)} = \frac{9}{21} \approx 0.43$
$b_2 = \frac{2(n^2+n+3}{9n(n-1)} = \frac{75}{504} \approx 0.148$
$c_1 = b_1 - \frac{1}{a_1} = \frac{9}{21} - \frac{1}{2.59} \approx 0.042$
$c_2 = b_2 - \frac{n+2}{a_1 n} + \frac{a_2}{a_1^2} \approx 0.04$
$e_1 = \frac{c_1}{a_1} = \frac{0.042}{2.59} \approx 0.016$
$e_2 = \frac{c_2}{a_1^2 + a_2} \approx 0.005$

$D = \frac{\Theta_\pi - \Theta_W}{\sqrt{e_1 S + e_2 S(S-1)}} = \frac{3.57 - 3.085}{\sqrt{0.016 \cdot 8 + 0.005 \cdot 8 \cdot 7}} \approx 0.76$

Task 3

This value suggests that this locus evolved in a population with a roughly constant size, since $-2 < D < 2$. There might be a light trend of a declining population, or the deviation from 0 is just due to the statistical error.

Task 4

Because Tajima's D of the African sample is relatively close to zero, this suggests that the population size in Africa was about constant during recent evolution, with at most small growth. Opposed to that, Tajima's D of the Asian sample is further in the negative, suggesting a growing population size in the course of this locus' evolution in the Asian region.

Task 5

The script calculated $\Theta_\pi - \Theta_W = 0.08$ with $n = 5, P = 20, S = 4$. Using the formula from Task 2 this results in:
$D = \frac{0.08}{0.29} = 0.27$