Two-allelic FST

Let

$$X_{ijk}, \quad i = 1, \ldots, m_k, j = 1, \ldots, n, k = 1, \ldots, r$$

be the allele at the j'th locus, in the ith individual in population k. Because we are in the two-allelic case $X_{ijk} \in \{0, 1\}$. Define

$$\begin{aligned}
m_\bullet &= m_1 + \cdots + m_r \\
f_{jk} &= \frac{1}{m_r} \sum_{i=1}^{m_r} X_{ijk} \\
f_j &= \sum_{k=1}^{r} \frac{m_k}{m_\bullet} f_{jk} \\
H_T^j &= 1 - f_j^2 - (1 - f_j)^2 \\
H_S^j &= \sum_{k=1}^{r} \frac{m_k}{m_\bullet} \big(1 - f_{jk}^2 - (1 - f_{jk})^2\big)
\end{aligned}$$

Then

$$F_{ST}^j = \frac{H_T^j - H_S^j}{H_T^j}$$

s-allelic FST

Let

$$X_{ijk}, \quad i = 1, \ldots, m_k, j = 1, \ldots, n, k = 1, \ldots, r$$

be the allele at the j'th locus, in the ith individual in population k. With s different alleles, $X_{ijk} \in \{A_1, \ldots, A_s\}$. Define

$$
\begin{aligned}
m_\bullet &= m_1 + \cdots + m_r \\
f_{jkp} &= \frac{1}{m_r} \sum_{i=1}^{m_r} 1(X_{ijk} = A_p) \\
f_{jp} &= \sum_{k=1}^{r} \frac{m_k}{m_\bullet} f_{jkp} \\
H_T^j &= 1 - \sum_{p=1}^{s} f_{jp}^2 \\
H_S^j &= \sum_{k=1}^{r} \frac{m_k}{m_\bullet} \left(1 - \sum_{p=1}^{s} f_{jkp}^2\right)
\end{aligned}
$$

Then

$$
\begin{aligned}
F_{ST}^j &= \frac{H_T^j - H_S^j}{H_T^j} \\
&= \frac{1 - \sum_{p=1}^{s} f_{jp}^2 - \sum_{k=1}^{r} \frac{m_k}{m_\bullet} \left(1 - \sum_{p=1}^{s} f_{jpk}^2\right)}{1 - \sum_{p=1}^{s} f_{jp}^2} \\
&= \frac{\sum_{k=1}^{r} \frac{m_k}{m_\bullet} \left(\sum_{p=1}^{s} f_{jpk}^2\right) - \sum_{p=1}^{s} \left(\sum_{k=1}^{r} \frac{m_k}{m_\bullet} f_{jkp}\right)^2}{1 - \sum_{p=1}^{s} f_{jp}^2} \\
&= \frac{\sum_{p=1}^{s} \left(\sum_{k=1}^{r} \frac{m_k}{m_\bullet} f_{jpk}^2 - \left(\sum_{k=1}^{r} \frac{m_k}{m_\bullet} f_{jkp}\right)^2\right)}{1 - \sum_{p=1}^{s} f_{jp}^2} \\
&= \frac{\sum_{p=1}^{s} \mathrm{Var}_{m_1,\ldots,m_r}\big((f_{j1p}, \ldots, f_{jrp})\big)}{1 - \sum_{p=1}^{s} f_{jp}^2}
\end{aligned}
$$

Until I have made this rather uninterpretable expression into something more interpretable, my only interpretation will be that F_{ST} is a quantity with historical roots that has proven to be useful in determining subpopulation structure and that it is related to something with the variance between populations(numerator) and something with the total variance(denominator).