Modeling Update

The statistical expression that has been used to model allelic drop-out by perturbing simulated mixture profiles should be updated.

On page 29 of “Characterization of Error Tradeoffs in Human Identity Comparisons: Determining a Complexity Threshold for DNA Mixture Interpretation,” allelic drop-out is modeled in the following manner:

\[
\Pr(D)\big|_{\phi}^{\alpha_{L,A}} = \frac{\Pr(D)\big|_{\phi=1}}{\phi}
\]

(Original Formulation)

\[\alpha_{L,A} \equiv \text{allele } A \text{ at locus } L\]
\[\phi \equiv \text{relative prevalence of allele (e.g., 0, 1, 2, 3, 4)}\]
\[\Pr(D)\big|_{\phi=1} \equiv \text{specified probability of drop-out for a heterozygous allele}\]
\[\Pr(D)\big|_{\phi}^{\alpha_{L,A}} \equiv \text{realized probability of drop-out for allele } A \text{ at locus } L\]

However, the relationship between an allele’s probability of drop-out given the binary detection scheme and relative prevalence model previously explicated in the thesis is more accurately captured with probability’s product rule. This results in a new formulation of the expression relating drop-out probability and relative prevalence:

\[
\Pr(D)\big|_{\phi}^{\alpha_{L,A}} = (\Pr(D))^{\phi}
\]

(New Formulation)

\[\alpha_{L,A} \equiv \text{allele } A \text{ at locus } L\]
\[\phi \equiv \text{relative prevalence of allele (e.g., 0, 1, 2, 3, 4)}\]
\[\Pr(D) \equiv \text{specified probability of drop-out for a heterozygous allele}\]
\[\Pr(D)\big|_{\phi}^{\alpha_{L,A}} \equiv \text{realized probability of drop-out for allele } A \text{ at locus } L \text{ with relative prevalence } \phi\]

Justification

Example 1: Consider an individual with heterozygous locus L. Adopting the notation employed above, the individual’s alleles at locus L are \(\alpha_{L,A1}\) and \(\alpha_{L,A2}\). Modeling allelic drop-out as a uniform, scalar quantity applied across all alleles at all loci, the probability of any particular allele dropping out is \(\Pr(D)\). Thus, the probability that \(\alpha_{L,A1}\) drops out is \(\Pr(D)\). Equivalently, the probability that \(\alpha_{L,A2}\) drops out is \(\Pr(D)\). Thus,

\[
\Pr(D)\big|_{\phi=1}^{\alpha_{L,A1}} = \Pr(D)\big|_{\phi=1}^{\alpha_{L,A2}} = \Pr(D) = (\Pr(D))^{1}.
\]
Example 2: Consider an individual with homozygous locus $L$. The individual’s alleles at locus $L$ are $a_{L,A1}$ and $a_{L,A1}$. Given a binary model of allele detection—in which detection occurs for all alleles with non-zero values for their relative prevalence—both instances of $a_{L,A1}$ must drop-out for $a_{L,A1}$ to have its relative prevalence reduced to zero, rendering it undetected and thus dropped-out. The probability of the first instance of $a_{L,A1}$ dropping out is $Pr(D)$, and equivalently, the probability of the second instance of $a_{L,A1}$ dropping out is $Pr(D)$. To calculate the probability of both events occurring—i.e., both instances of $a_{L,A1}$ dropping out—the product rule must be invoked. The probability of $Pr(D)$ AND $Pr(D)$ is the product $Pr(D) \times Pr(D)$. Thus, $Pr(D)^{\phi_{a_{L,A1}}=2} = Pr(D) \cdot Pr(D) = (Pr(D))^2$.

Example 3: Consider a two-person mixture comprised of individuals with overlapping alleles at locus $L$, one of whom is homozygous at $L$ while the other is a heterozygote. The discrete mixture alleles at $L$ contributed from the individuals might be $a_{L,A1}$, $a_{L,A1}$, $a_{L,A1}$, and $a_{L,A2}$. In this case, for $a_{L,A1}$ to drop-out, the first instance AND the second instance AND the third instance of $a_{L,A1}$ must all drop-out for $a_{L,A1}$ to go undetected. Thus,

$$Pr(D)^{\phi_{a_{L,A1}}=3} = Pr(D) \cdot Pr(D) \cdot Pr(D) = (Pr(D))^3.$$  

Example 4: Consider a two-person mixture comprised of individuals with overlapping, homozygous alleles at locus $L$. The mixture’s alleles at $L$ are $a_{L,A1}$, $a_{L,A1}$, $a_{L,A1}$, and $a_{L,A1}$. In order for $a_{L,A1}$ to go undetected, all four instances of the allele must drop-out. In this case,

$$Pr(D)^{\phi_{a_{L,A1}}=4} = Pr(D) \cdot Pr(D) \cdot Pr(D) \cdot Pr(D) = (Pr(D))^4.$$

The proper relationship between the probability of drop-out of a particular allele $A$ at locus $L$, given by $a_{L,A}$, with relative prevalence $\phi$ can be inductively reasoned to be $Pr(D)^{\phi_{a_{L,A}}=\phi} = (Pr(D))^\phi$.

**Implications for Results**

The updated expression relating the probability of a given allele’s drop-out to the power of a heterozygous allele’s probability of drop-out should be employed to re-synthesize the simulated mixture profiles contained in the experiment. Once a new collection of simulated mixture profiles for levels of drop-out (corresponding to a heterozygous allele) of 0.10 through 0.90 in increments of 0.10 is completed, comparisons between these new simulated mixtures and the existing populations of simulated excluded individuals and simulated included individuals, respectively, must be redone.

The existing populations of simulated excluded individuals and simulated included individuals is unaffected by this updated drop-out modeling. Similarly, the laboratory mixtures as well as their associated simulated included individuals are unaffected.