8 April 2025 · CC-BY 4.0 Qeios

Peer Review

Review of: "Misestimation of Expected Genetic Differences – A Statistical Note on Some Recent Papers"

Richárd Kőhegyi[1]

1. Independent researcher

When discussing Gusev, Furest makes several errors:

(1): Conflates $(\frac{\sigma^2_{AB}}{\sigma^2_{POP}})$ with (σ^2_{AB})

(2): Possibly due to (1), Fuerst misunderstands and therefore misstates the origin of the "2" multiplier.

(3) Misrepresents the paper by Edge&Rosenberg. They state "Thus, regardless of ploidy ℓ, Q_{ST} is obtained from the expression in Eq. 27 by setting ℓ to 1."(when h^2=1). They also state "that is, if $QST = FkST$, then $(\rho^2)T = FkST(\ell)$."

In the following derivations, I attempt to clarify the misunderstandings behind (1) and (2):

First, note that when Piffer quotes Gusev: "As a consequence, "*the proportion of heritable variance in the trait attributable to genetic differences between the populations*" (Q_{ST}) is approximately equal to the cross-population (Hudson) F_{ST} (which we rederive as $F_{ST,1}$). This quantity also does not depend on polygenicity. *Note*: the relationship is sometimes reported as 2*F_{ST}*h2, but this is only an approximation for Nei's F_{ST}: Nei's F_{ST} is approximately equal to half of Hudson's F_{ST} when the former is close to 0 or 1 (see [8.6]).",

Gusev refers to $(\frac{\sigma^2_{AB}}{\sigma^2_{POP}})$*h^2, NOT (σ^2_{AB})*h^2 as Fuerst seems to(?) assume

In the derivations, I assume neutrality and that FST is small:

We start from the standard diploid Q_{ST} definition:

$$Q_{ST} = \frac{\sigma^2_{AB}}{\sigma^2_{AB} + 2\sigma^2_{AW}},$$

and, assuming neutrality, we set

$$Q_{ST} \approx F_{ST}.$$

Case 1: Using Hudson's F_{ST}

Hudson's estimator is given by

$$F_{ST,\text{Hudson}} \approx \frac{(p_1 - p_2)^2}{p_1 q_2 + p_2 q_1},$$

with $q_i = 1 - p_i$. Note that

$$p_1 q_2 + p_2 q_1 = p_1(1 - p_2) + p_2(1 - p_1) = p_1 + p_2 - 2p_1 p_2,$$

and

$$(p_1 - p_2)^2 = p_1^2 - 2p_1 p_2 + p_2^2.$$

Substitute into the neutrality assumption:

$$\frac{\sigma^2_{AB}}{\sigma^2_{AB} + 2\sigma^2_{AW}} \approx \frac{(p_1 - p_2)^2}{p_1 q_2 + p_2 q_1},$$

and cross-multiply:

$$\sigma^2_{AB}\,(p_1 + p_2 - 2p_1 p_2) \approx (\sigma^2_{AB} + 2\sigma^2_{AW})\,(p_1^2 - 2p_1 p_2 + p_2^2).$$

Expanding the right-hand side gives

$$\sigma^2_{AB}(p_1 + p_2 - 2p_1 p_2) \approx \sigma^2_{AB}(p_1^2 - 2p_1 p_2 + p_2^2) + 2\sigma^2_{AW}(p_1^2 - 2p_1 p_2 + p_2^2).$$

Writing the left-hand side explicitly:

$$p_1\sigma^2_{AB} + p_2\sigma^2_{AB} - 2p_1 p_2\sigma^2_{AB} \approx p_1^2\sigma^2_{AB} - 2p_1 p_2\sigma^2_{AB} + p_2^2\sigma^2_{AB} + 2(p_1 - p_2)^2\sigma^2_{AW}$$

Cancel the common $-2p_1 p_2\sigma^2_{AB}$ terms:

$$p_1\sigma^2_{AB} + p_2\sigma^2_{AB} \approx p_1^2\sigma^2_{AB} + p_2^2\sigma^2_{AB} + 2(p_1 - p_2)^2\sigma^2_{AW}.$$

Rearrange by gathering the σ^2_{AB} terms:

$$\left[p_1 + p_2 - (p_1^2 + p_2^2)\right]\sigma^2_{AB} \approx 2(p_1 - p_2)^2\sigma^2_{AW}.$$

Since

$$p_1 + p_2 - p_1^2 - p_2^2 = p_1(1 - p_1) + p_2(1 - p_2) = p_1 q_1 + p_2 q_2,$$

we obtain

$$(p_1 q_1 + p_2 q_2)\sigma^2_{AB} \approx 2(p_1 - p_2)^2\sigma^2_{AW}.$$

Solving for σ^2_{AB}:

$$\boxed{\sigma^2_{AB} \approx \frac{2(p_1 - p_2)^2}{p_1 q_1 + p_2 q_2}\,\sigma^2_{AW} \quad \text{(Using Hudson's } F_{ST}\text{)}.}$$

doi.org/10.32388/F1H4MP

Case 2: Using Nei's 1973 F_{ST}

Nei's estimator is given by

$$F_{ST,\text{Nei73}} \approx \frac{(p_1 - p_2)^2}{4p_T q_T},$$

where p_T is the total allele frequency (often taken as $p_T = \frac{p_1+p_2}{2}$) and $q_T = 1 - p_T$. Notice that

$$4p_T q_T = 4\left(\frac{p_1+p_2}{2}\right)\left(1 - \frac{p_1+p_2}{2}\right) = (p_1+p_2)\Big(2-(p_1+p_2)\Big) = 2(p_1+p_2) - (p_1+p_2)^2.$$

Substitute into the neutrality assumption:

$$\frac{\sigma^2_{AB}}{\sigma^2_{AB} + 2\sigma^2_{AW}} \approx \frac{(p_1-p_2)^2}{4p_T q_T},$$

and cross-multiply:

$$\sigma^2_{AB}\Big[2(p_1+p_2) - (p_1+p_2)^2\Big] \approx (\sigma^2_{AB} + 2\sigma^2_{AW})(p_1-p_2)^2.$$

Expanding the left-hand side:

$$\sigma^2_{AB}\Big(2p_1 + 2p_2 - p_1^2 - 2p_1p_2 - p_2^2\Big) \approx \sigma^2_{AB}(p_1^2 - 2p_1p_2 + p_2^2) + 2\sigma^2_{AW}(p_1-p_2)^2.$$

Cancel the common $-2p_1p_2\sigma^2_{AB}$ terms:

$$2p_1\sigma^2_{AB} + 2p_2\sigma^2_{AB} - p_1^2\sigma^2_{AB} - p_2^2\sigma^2_{AB} \approx p_1^2\sigma^2_{AB} + p_2^2\sigma^2_{AB} + 2(p_1-p_2)^2\sigma^2_{AW}.$$

Rearrange to bring like terms together:

$$\Big[2p_1 + 2p_2 - 2p_1^2 - 2p_2^2\Big]\sigma^2_{AB} \approx 2(p_1-p_2)^2\sigma^2_{AW}.$$

Dividing by 2:

$$(p_1 + p_2 - p_1^2 - p_2^2)\sigma^2_{AB} \approx (p_1-p_2)^2\sigma^2_{AW}.$$

Since

$$p_1 + p_2 - p_1^2 - p_2^2 = p_1q_1 + p_2q_2,$$

we obtain

$$(p_1q_1 + p_2q_2)\sigma^2_{AB} \approx (p_1-p_2)^2\sigma^2_{AW}.$$

Thus, solving for σ^2_{AB}:

$$\boxed{\sigma^2_{AB} \approx \frac{(p_1-p_2)^2}{p_1q_1 + p_2q_2}\,\sigma^2_{AW} \quad \text{(Using Nei's 1973 } F_{ST}\text{).}}$$

Summary and Comparison

The two derivations yield:

$$\text{Hudson:}\quad \sigma^2_{AB} \approx \frac{2(p_1-p_2)^2}{p_1q_1 + p_2q_2}\,\sigma^2_{AW},$$

$$\text{Nei (1973):}\quad \sigma^2_{AB} \approx \frac{(p_1-p_2)^2}{p_1q_1 + p_2q_2}\,\sigma^2_{AW}.$$

Hopefully, one can easily see that it does matter which FST we plug into the QST equation! The (σ^2_{AB}) is 2x larger when we use Hudson's FST compared to Nei's FST

Next, I show how the(σ^2_{AB}) relates to$(\frac{\sigma^2_{AB}}{\sigma^2_{POP}})$

Define $h^2 = (\sigma^2_{AB} + 2\sigma^2_{AW})/\sigma^2_P$. This implies $\frac{\sigma^2_{AB}}{\sigma^2_P} = h^2 Q_{ST}$, where $Q_{ST} = \frac{\sigma^2_{AB}}{\sigma^2_{AB}+2\sigma^2_{AW}}$.

The derivations yielded:

$$\sigma^2_{AB(H)} \approx \frac{2(p_1 - p_2)^2}{p_1q_1 + p_2q_2}\sigma^2_{AW} \tag{1}$$

$$\sigma^2_{AB(N)} \approx \frac{(p_1 - p_2)^2}{p_1q_1 + p_2q_2}\sigma^2_{AW} \tag{2}$$

Note $\sigma^2_{AB(H)} \approx 2\sigma^2_{AB(N)}$.

Substituting (1) into the h^2Q_{ST} expression and applying small F_{ST} approximations ($p_1q_1 + p_2q_2 \approx p_1q_2 + p_2q_1$; denominator $1 + \cdots \approx 1$):

$$\frac{\sigma^2_{AB(H)}}{\sigma^2_P} = h^2\left(\frac{\sigma^2_{AB(H)}}{\sigma^2_{AB(H)} + 2\sigma^2_{AW}}\right) \approx h^2\left(\frac{(p_1 - p_2)^2}{p_1q_2 + p_2q_1}\right) \approx h^2F_{ST,Hudson}$$

$$\boxed{\frac{\sigma^2_{AB}}{\sigma^2_P} \approx h^2F_{ST,Hudson}}$$

Substituting (2) into the h^2Q_{ST} expression and applying small F_{ST} approximations ($2(p_1q_1+p_2q_2) \approx 4p_Tq_T$; denominator $1 + \cdots \approx 1$):

$$\frac{\sigma^2_{AB(N)}}{\sigma^2_P} = h^2\left(\frac{\sigma^2_{AB(N)}}{\sigma^2_{AB(N)} + 2\sigma^2_{AW}}\right) \approx h^2\left(\frac{(p_1 - p_2)^2}{4p_Tq_T}\right) \approx h^2F_{ST,Nei73}$$

Since $\sigma^2_{AB(H)} \approx 2\sigma^2_{AB(N)}$, we have $\frac{\sigma^2_{AB(H)}}{\sigma^2_P} \approx 2\frac{\sigma^2_{AB(N)}}{\sigma^2_P}$. Therefore, using the standard Hudson-based result:

$$\frac{\sigma^2_{AB}}{\sigma^2_P} \approx h^2F_{ST,Hudson} \approx 2(h^2F_{ST,Nei73})$$

$$\boxed{\frac{\sigma^2_{AB}}{\sigma^2_P} \approx 2h^2F_{ST,Nei73}}$$

As we can see, Gusev's claim was correct. I hope this review clears some misunderstandings up.

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/F1H4MP