

## POPULATION GENETICS PROBLEM #4

The data for this problem consist of half-sib families.<sup>1</sup> That means there are only two components of variance that we can estimate: the variance within sibships,  $\sigma_w^2$ , and the variance among sibships,  $\sigma_d^2$ . Here's the code that we need to do this, and to estimate the causal components of variance and the heritability.

```
model {
  # dams
  for (i in 1:23) {
    beta[i] ~ dnorm(0.0, tau.dam)
  }
  beta[24] <- -sum(beta[1:23])
  # offspring
  for (i in 1:190) {
    mu[i] <- nu + beta[dam[i]] + gamma*temp[i] + delta*age[i]
    trunkv[i] ~ dnorm(mu[i], tau.within)
  }

  # observational components
  v.dam <- 1.0/tau.dam
  v.within <- 1.0/tau.within
  v.p <- v.dam + v.within

  # causal components
  v.a <- 4*v.dam
  h.2 <- v.a/v.p

  # priors
  nu ~ dnorm(0.0, 0.001)
```

---

<sup>1</sup>More precisely they consist of families we *assume* are half sibs. We don't actually know that they are half sibs.

```

gamma ~ dnorm(0.0, 0.001)
delta ~ dnorm(0.0, 0.001)
tau.within ~ dgamma(1, 0.25)
k <- 3*tau.within
tau.dam ~ dgamma(1, 0.25)I(k,)
}

```

And here are the results:

The screenshot shows a window titled 'Log' containing a table of MCMC node statistics. The table is organized into seven sections, each starting with 'Node statistics' and followed by a header row: **node**, **mean**, **sd**, **MC error**, **2.5%**, **median**, **97.5%**, **start**, and **sample**. The data rows are as follows:

Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
v.dam	0.07135	0.01593	1.927E-4	0.04074	0.07167	0.1017	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
v.within	0.2955	0.03455	3.473E-4	0.2369	0.2925	0.3695	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
v.p	0.3668	0.03949	3.492E-4	0.2987	0.364	0.4516	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
v.a	0.2854	0.06374	7.71E-4	0.1629	0.2867	0.4067	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
h.2	0.7778	0.1491	0.001982	0.4561	0.7958	0.9912	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
gamma	0.1917	0.07824	9.176E-4	0.03558	0.1932	0.3418	10001	10000
Node statistics								
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
delta	0.02324	0.08025	9.933E-4	-0.1342	0.02291	0.185	10001	10000

1. What is the value of the phenotypic variance in number of trunk vertebrae,  $V_p$ , in this population?

The phenotypic variance is simply the sum of the within and among dam components of variance, i.e.,  $v.p$ . So the phenotypic variance in number of trunk vertebrae is about 0.37, with a 95% credible interval of (0.30, 0.45).

2. What is the value of the additive genetic variance in number of trunk vertebrae,  $V_a$ , in this population?

Since we're assuming that the offspring of dams are half sibs, the among sibship component of variance is equal to  $\frac{1}{4}V_a$ , so  $V_a = 4\sigma_d^2$ , i.e., v. a. The additive genetic variance in trunk vertebrae is about 0.29, with a 95% credible interval of (0.16, 0.41).

3. What is the heritability of number of trunk vertebrae in this population?

I didn't specify that I wanted the *narrow sense* heritability, but since you can only estimate  $V_a$  from a half-sib breeding design, you can only estimate  $h_N^2 = \frac{V_a}{V_p}$ , i.e., h.2. The heritability of trunk vertebrae number is about 0.78, with a 95% credible interval of (0.46, 0.99).<sup>2</sup>

4. Is there evidence that developmental temperature influences the number of trunk vertebrae in this population?

The influence of developmental temperature is reflected in the coefficient **gamma**. If **gamma** is positive, it means that increasing temperatures are associated with increasing numbers of vertebrae. If it is negative, it means that increasing temperatures are associated with decreasing numbers of vertebrae. The estimate of **gamma** 0.19 with a 95% credible interval (0.036, 0.34) that does not overlap zero, indicating that higher developmental temperatures are associated with more trunk vertebrae.

5. Is there evidence that the number of trunk vertebrae depends on the age at which an individual is measured?

The influence of age at measurement is reflected in the coefficient **delta**. Since the 95% credible intervals broadly overlap zero, we have no evidence that the age at which animals are measured influences the number of vertebrae that are counted.<sup>3</sup>

6. **Warning: This one's hard.** Your analysis assumed that all offspring were half-sibs, i.e., that they all had different fathers. Would your estimates of  $V_a$  and  $h_N^2$  be larger, smaller, or unchanged if some of the offspring were full sibs? Why?

---

<sup>2</sup>In case your wondering part of the reason I chanted the magic for the priors the way I did was to constrain the heritability estimate to be less than 1. Just for the heck of it, you might want to retry the analysis with out the  $I(\mathbf{k},)$  on the prior for **tau.dam** and see what you get.

<sup>3</sup>That's probably not surprising, but it's nice to see statistics confirming what you probably would have guessed anyway.

In estimating  $V_a$  we assumed that the progeny are all half sibs, so that  $\sigma_d^2 = \frac{1}{4}V_a$ . Suppose, on the contrary, that all of the progeny were full sibs. Then

$$\begin{aligned}\sigma_d^2 &= \frac{1}{2}V_a + \frac{1}{4}V_d \\ &> \frac{1}{4}V_a \quad .\end{aligned}$$

Thus, we *overestimate*  $V_a$  by assuming that all progeny are half sibs, by at least a factor of two. In practice,  $V_d$  is often small, so our estimates might not be too far off if we estimated  $V_a$  as  $2\sigma_d^2 = 0.1427$ , rather than  $4\sigma_d^2 = 0.2854$ . Thus, if some of the offspring were full sibs, we'd be overestimating  $V_a$  (and  $h_N^2$ ).

I should also point out that I didn't tell the whole truth in the equation above. If there are non-genetic maternal effects, larger females producing offspring with more vertebrae for example, or common environment effects, both will be included in  $\sigma_d^2$ , leading to an even greater overestimation of  $V_a$ .