Heritability

James Heckman
University of Chicago

Econ 350, Winter 2023

Twin Methods

- *Basic principle.* Monozygotic (identical) twins are more similar than dizygotic (fraternal) twins.

- *Key assumption.* If environmental similarities are the *same* for both types of twins, then we can estimate genetic components of outcomes.

Univariate Twin Model

$$Y = \text{ observed “phenotypic” variable}$$

$$X = \text{ unobserved “genotype”}$$

$$U = \text{ environment}$$

- Assume additivity:

$$Y = X + U$$

- This assumption is critical to the literature, but is not justified (see, e.g. Rutter, Caspi and Moffitt, 2006; Rutter, 2006).

- Assume independence:

$$\sigma_Y^2 = \sigma_X^2 + \sigma_U^2$$

- Pair each person with another individual:

$$Y^* = X^* + U^*$$

- Phenotypic covariance:

$$\mathrm{Cov}(Y, Y^*) = \mathrm{Cov}(X, X^*) + \mathrm{Cov}(U, U^*)$$

- Assume

$$\mathrm{Cov}(X, U^*) = \mathrm{Cov}(X^*, U) = 0.$$

- Use standardized form:

$$x = \frac{X}{\sigma_X} \qquad u = \frac{U}{\sigma_U} \qquad y = \frac{Y}{\sigma_Y}$$

$$h^2 = \frac{\sigma_X^2}{\sigma_Y^2} \qquad e^2 = \frac{\sigma_U^2}{\sigma_Y^2} \qquad y = hx + eu$$

- Therefore $h^2 + e^2 = 1$ and

$$C = \text{Correl}(Y, Y^*) = gh^2 + \rho e^2, \text{ where}$$

$$g = \frac{\text{Cov}(X, X^*)}{\text{Cov}(X, X)}$$

$$\rho = \frac{\text{Cov}(U, U^*)}{\text{Var}(U)}$$

- For monozygotes and for dizygotes,

$$g_{MZ} = 1 \qquad g_{DZ} < 1$$

$$C_{MZ} = h^2 + \rho_{MZ} e^2$$

$$C_{DZ} = g_{DZ} h^2 + \rho_{DZ} e^2$$

- $$C_{MZ} - C_{DZ} = (1 - g_{DZ}) h^2 + (\rho_{MZ} - \rho_{DZ}) e^2$$

- Using the identity $h^2 + e^2 = 1$,

$$h^2 = \frac{(C_{MZ} - C_{DZ}) - (\rho_{MZ} - \rho_{DZ})}{(1 - g_{DZ}) - (\rho_{MZ} - \rho_{DZ})}.$$

- One equation for h in two unknowns:

  1. $(\rho_{MZ} - \rho_{DZ})$
  2. $(1 - g_{DZ})$

- Jensen assumes that

$$\rho_{MZ} = \rho_{DZ},$$

- Therefore

$$h^2 = \frac{C_{MZ} - C_{DZ}}{1 - g_{DZ}}.$$

- $$g_{DZ} = 1/2 \qquad \text{random mating}$$

  $$g_{DZ} = 2/3 \qquad \text{strong assortive mating}$$

- *Issue.* Might the *environment* treat MZ (identical) twins more alike than DZ?
- Taubman: socioeconomic status (income, occupation) has $C_{MZ} = 0.6$, $C_{DZ} = 0.4$.
- Jensen IQ: $C_{MZ} = 0.87$, $C_{DZ} = 0.56$.
- Suppose

$$g_{DZ} = 1/2 \quad C_{MZ} - C_{DZ} = 0.2 \quad \rho_{MZ} - \rho_{DZ} = 0.2$$

$$h^2 = \frac{0.2 - 0.2}{(1 - 1/2) - 0.2} = 0.$$

- Others gild the lily by assuming

$$\mathrm{Cov}(X, U) \neq 0.$$

- A fundamental identification problem.
- Other assumptions are also made.
- Issues in multivariate settings:
  1. linearity
  2. interpretation

Heritability Estimates from Martin (1975)

Subject	Estimated h^2
English	0.79 ± 0.05
French	0.83 ± 0.07
History	0.47 ± 0.13
Geography	0.81 ± 0.06
Mathematics 1	0.81 ± 0.05
Mathematics 2	0.81 ± 0.06
Physics	0.77 ± 0.09
Chemistry	0.89 ± 0.05
Science 1	0.75 ± 0.09
Science 2	0.77 ± 0.08
IQ	0.79 ± 0.06

Univariate Heritability Estimates from Behrman, Taubman and Wales Data

	Schooling	Initial Occupation	Current Occupation	Log Earnings
Observed Δc	0.22	0.20	0.23	0.24
Estimated h^2				
Assuming $\Delta g = 2/3$:	0.33	0.30	0.34	0.36
Assuming $\Delta g = 1/2$:	0.44	0.40	0.46	0.48
Assuming $\Delta g = 1/3$:	0.66	0.60	0.69	0.72