Measures of Assortativity

The Toulouse Economics and Biology Workshop

Ted Bergstrom and Yuji Tamira

University of California Santa Barbara and Australian National University

May 19, 2013

What is Assortativity?

- Suppose that a population has two or more types of individuals who form into subgroups, as determined by some assignment process.
- The degree of assortativity of an assignment process is a measure of the propensity of "likes to be matched with likes".
- This propensity depends, in general, both on the proportions of types in the population and on the nature of the process.
- We will explore alternative matching processes and measures of assortativity of these processes.

Example: Two-Pool Assortative Process

- There are two types.
- With probability F an individual selects a partner from an *assortative pool* that includes only members of this individual's type.
- With probability $1 - F$, an individual selects a partner at random from a *random pool* that includes everyone who did not join an assortative pool.
- Let's call F the *assortativity* of this process.

Measurement of assortativity

- If we observe the composition of matched pairs resulting from a two-pool assortative process, how could we estimate the assortativity F?
- Suppose we observe a population in which the proportion of Type 1's is $\hat{p}$ and the proportion of Type 2's is $1 - \hat{p}$. Let $\hat{\pi}_{ij}$ be the observed fraction of all pairs that have one member of type i and one of type j.
- Then we can estimate F by

$$1 - \frac{\hat{\pi}_{12}}{2\hat{p}(1 - \hat{p})}.$$

Proof of this claim.

- The probability π_{12} that a randomly chosen pair has one member of each type is the probability that a randomly chosen individual from the population is either a type 1 who is matched with a type 2, or a type 2 matched with a type 1.
- The probability that a randomly chosen individual is of type 1 who is matched to a type 2 is the probability that a random draw is a type 1 who joins the random pool and happens to draw a type 2 for a partner. This probability is

$$p(1-F)(1-p)$$

.

- The probability that a randomly chosen individual is a type 2 matched with a type 1 is

$$(1-p)(1-F)p = p(1-F)(1-p).$$

Proof continued:

- Therefore
$$\pi_{12} = 2p(1-p)(1-F).$$
- Rearranging terms, we have
$$F = 1 - \frac{\pi_{12}}{2p(1-p)}.$$
Since $\hat{p}$ and $\hat{\pi}_{12}$ are maximum likelihood estimates of p and π_{12}, we estimate F as
$$\hat{F} = 1 - \frac{\hat{\pi}_{12}}{2\hat{p}(1-\hat{p})}.$$

Equivalent Measures in Biology:

- *Coefficient of Inbreeding:* "the correlation between homologous genes of uniting gametes under a given mating pattern."
  - (Sewall Wright's F-Statistic.) Wright applied this measure to inbreeding of kin by animal breeders as well as to preferential mating patterns based on phenotypic similarities
- *Coefficient of Relatedness* between two individuals: "the probability that a rare gene possessed by one is also possessed by the other."
  - Foundation of Hamilton's theory of kin selection.
- *Reduction in Heterozygosity* "the fractional reduction in heterozygosity relative to a random-mating population with the same allele frequencies." (Hartl and Clark, 1959).

One more equivalent measure

- *Index of Assortativity*: The difference between the probability of being matched with a given type if one is also of this type and the probability of being matched with that type if one is of another type.
  - Useful way of thinking about evolutionary dynamics of random matching processes.

Genetic relatedness and the two-pool assortative model

- The two-pool assortative process models assortative matching of genes, where two individuals might inherit the same gene in a given locus for one of two reasons.
  - They inherited this gene from a common ancestor. (They draw their partner from the same assortative pool)
  - They did not inherit the gene from a common ancestor, but happened to draw the same gene anyway. (They draw their partner the random pool).
  - In many structured environments, the probability that two individuals e.g. two siblings, two cousins, inherit the same gene in a given locus is constant and independent of proportions of the genes in the population.

Application to Evolutionary Dynamics: Hamilton's Helping Game

- Players are matched in pairs. Each player can exert a level of effort x to help the other. Where x_i is the effort level of player i, the reproductive fitness of player 1 is $b(x_2) - c(x_1)$ and that of player 2 is $b(x_1) - c(x_2)$.
- Assume b is an increasing, concave function and c an increasing convex function.
- Each player has its own type, which is the amount of help x that it will offer.
- Matching is assortative according to a two-pool process. With probability F, a player finds its match in an assortative pool consisting only of its own type. With probability $1 - F$ it is matched with a random draw from those who were not assigned to assortative pools.

- If the reproduction rate of each type is increasing in its fitness, then there is a unique "evolutionary equilibrium" in which all individuals are of a type x^* such that x^* maximizes $b(x)F - c(x)$.
- Consider a population with some type x^* individuals and some of other types. Let y be the expected value of $b(x)$ that someone would receive from a random assignment.
- Expected payoff to a type x is $b(x)F + y(1 - F) - c(x)$.
- Therefore If x^* maximizes $b(x)F - c(x)$, the x^* type will have a higher payoff than any other.

Variable assortativity.

- For a population with proportions p and $1-p$ of types 1 and 2 in which the fraction π_{12} of pairs include one person of each type, define
$$F(p) = 1 - \frac{\pi_{12}}{2p(1-p)}.$$
- For the two-pool assortative process, $F(p) = F$ and is independent of p.
- The genetic processes studied by Wright and Hamilton also have $F(p)$ independent of p and are equivalent to the two-pool process.
- As we will show, there are some interesting matching processes for which $F(p)$ is not independent of p.

Empirical Example: Black-White Marriages

- U.S. Census ACS survey of married couples reports race of each member of a random sample of married couples in each state.
- Suppose that matching in each state is by a two-pool assortative process with uniform assortativity.
- Let $\hat{p}_i$ be the fraction of married individuals in the sample for state i who are black and $1 - \hat{p}_i$ the fraction who are not.
- Let π_{BWi} is the fraction of observed marriages in state i that include one person of each race.
- Then estimate $\hat{F}_i$ for state i by

$$\hat{F}_i = 1 - \frac{\hat{\pi}_{BWi}}{2\hat{p}_i(1 - \hat{p}_i)}.$$

Sample results

Table: Asssortativity for Selected States

State	Fraction of Blacks Married to Whites	Fraction of Whites Married to Blacks	$\hat{F}$
Arizona	0.226	.0057	0.77
California	0.136	.0070	0.86
New York	0.059	.0065	0.93
Mississippi	0.015	.0043	0.98

Regression Results

- We regressed the state estimates $\hat{F}_i$ on 5 variables, all of which had statistically effects on assortativity. The regression "explained" about 75% of variation.
  - Former slave state with anti-miscegenation laws in force in 1967. (positive)
  - Index of residential segregation (measures differences in racial composition of census tracts) (positive).
  - Percent living in metropolitan areas (positive).
  - Percent of population who are black (positive).
  - Percent of population with college degrees (negative).
- Significance of percent of population who are black suggests that the simple two pool theory may not be adequate.

- Since evidence suggests the

$$\frac{\pi_{BW}}{2p(1-p)}$$

is not independent of p, it is worthwhile to consider alternative assortment theories.
- We consider two such processes.
    - A non-uniform two pool assortative process in which one type is more likely to seek a partner in an assortative pool than the other.
    - A "strangers-in-the-dark" process where meetings occur randomly, but the probability that they a meeting leads to a relationship is higher for two individuals of the same type than for two of different types.

A Non-uniform two-pool process

- Suppose that one of the two types is more likely to join the assortative pool than the other.
- Let F_i be the probability that a type i joins the assortative pool for $i = 1, 2$ and let $G_i = 1 - F_i$ be the probability that a type i joins the random pool.
- The probability that a couple consists of of one person of each type is

$$\pi_{12} = 2p(1-p)\frac{G_1 G_2}{pG_1 + (1-p)G_2}.$$

- With this non-uniform two-pool process, we have

$$F(p) = 1 - \frac{\pi_{12}}{2p(1-p)} = 1 - \frac{G_1 G_2}{pG_1 + (1-p)G_2},$$

  which is not independent of p.
- In this case, if $G_1 > G_2$, then $F(p)$ is an increasing function of p.

Estimating non-uniform two-pool marriage process with U.S. states

- Preliminary estimates suggest that $G_W > G_B$, which means that whites are more likely to seek matches in segregated assortative pools than blacks.
- This is consistent with our regressions showing that F is higher in states with a larger percentage of blacks.

Strangers-in-the-Night Assortative Processes

- Pairs of individuals meet randomly. If they are of the same type, they form a pair with probability s. If they are of different types, they form a pair with probability $m < s$.
- Where p is the fraction of type 1's and $q = 1 - p$, the expected ratios of partnership types in the population are:

$$\pi_{11} = \frac{p^2 s}{(p^2 + q^2) + 2pqm}$$

$$\pi_{22} = \frac{q^2 s}{(p^2 + q^2) + 2pqm}$$

$$\pi_{12} = \frac{2pqm}{(p^2 + q^2) + 2pqm}$$

Assortativity of Strangers-in-the-Night

- For a strangers-in-the-night process where individuals who meet will pair up with probability s if they are the same type and probability m if they are of different types, the degree of assortativity is

$$F(p) = 1 - \frac{\pi_{12}}{2p(1-p)} = \frac{pq(s^2 - m^2)}{pq(s^2 - m^2) + sm}.$$

- Where $s > m$, $F(p)$ is increasing in the size of the minority population and is maximized when $p = 1/2$.
- This is also consistent with our comparisons of the assortativity of black-white marriages in the U.S.

Estimating Strangers-in-the-Night parameters for Black-white marriages

- Where $\bar{\pi}_{ij}$ is the expected value of the random variable π_{ij}, we have
$$\frac{m}{s} = \frac{\bar{\pi}_{BW}}{2\sqrt{\bar{\pi}_{BB}\bar{\pi}_{WW}}}$$
- Thus we estimate the ratio $\frac{m}{s}$ as
$$\left(\frac{\hat{m}}{s}\right) = \frac{\hat{\pi}_{BW}}{2\sqrt{\hat{\pi}_{BB}\hat{\pi}_{WW}}}$$

Sample results

Table: Asssortativity Measures

State	Fraction of Blacks Married to Whites	Fraction of Whites Married to Blacks	$\hat{F}$	Ratio m/s
Arizona	0.226	.0057	0.77	0.55
California	0.136	.0070	0.86	0.41
New York	0.059	.0065	0.93	0.26
Mississippi	0.015	.0043	0.98	0.14