Odds ratios for covariates effects

567 Statistical analysis of social networks

Peter Hoff

Statistics, University of Washington

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- Covariate: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- Covariate: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- Covariate: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Statistics for covariate effects

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

Often we also have node-level **covariate** information.

- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities

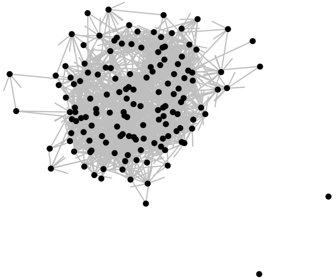
Often we also have node-level **covariate** information.

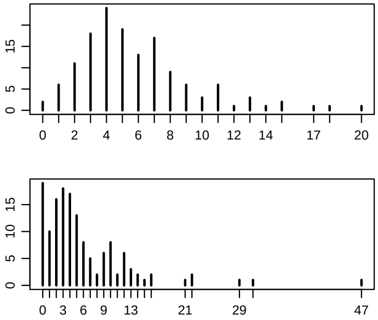
- **Covariate**: Node characteristics that "co-vary" with the network.

Questions:

- How to describe the relationship between the network and covariates?
- Can the covariates explain/predict network behavior?

Example: Girls friendships






```
mean(Y,na.rm=TRUE)

## [1] 0.04088967

Ce( 1*( Y+t(Y) > 0 )  )

## [1] 0.3820283
```

Covariate effects

We also have data on GPA

- `hgpa` = indicator of above-average gpa;

```
mean( Y[ hgpa==1, hgpa==1] ,na.rm=TRUE)

## [1] 0.04737443

mean( Y[ hgpa==1, hgpa==0] ,na.rm=TRUE)

## [1] 0.037623

mean( Y[ hgpa==0, hgpa==1] ,na.rm=TRUE)

## [1] 0.03935944

mean( Y[ hgpa==0, hgpa==0] ,na.rm=TRUE)

## [1] 0.03903421
```

Covariate effects

We also have data on GPA

- `hgpa` = indicator of above-average gpa;

```
mean( Y[ hgpa==1, hgpa==1] ,na.rm=TRUE)

## [1] 0.04737443

mean( Y[ hgpa==1, hgpa==0] ,na.rm=TRUE)

## [1] 0.037623

mean( Y[ hgpa==0, hgpa==1] ,na.rm=TRUE)

## [1] 0.03935944

mean( Y[ hgpa==0, hgpa==0] ,na.rm=TRUE)

## [1] 0.03903421
```

Covariate effects

We also have data on GPA

- hgpa = indicator of above-average gpa;

```
mean( Y[ hgpa==1, hgpa==1] ,na.rm=TRUE)

## [1] 0.04737443

mean( Y[ hgpa==1, hgpa==0] ,na.rm=TRUE)

## [1] 0.037623

mean( Y[ hgpa==0, hgpa==1] ,na.rm=TRUE)

## [1] 0.03935944

mean( Y[ hgpa==0, hgpa==0] ,na.rm=TRUE)

## [1] 0.03903421
```

Covariate effects

We also have data on smoking behavior:

- `hsmoke` = indicator of above-average smoking behavior.

```
mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04477612

mean( Y[ hsmoke==1, hsmoke==0] ,na.rm=TRUE)

## [1] 0.03062609

mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04400078

mean( Y[ hsmoke==0, hsmoke==0] ,na.rm=TRUE)

## [1] 0.04425837
```

Covariate effects

We also have data on smoking behavior:

- `hsmoke` = indicator of above-average smoking behavior.

```
mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04477612

mean( Y[ hsmoke==1, hsmoke==0] ,na.rm=TRUE)

## [1] 0.03062609

mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04400078

mean( Y[ hsmoke==0, hsmoke==0] ,na.rm=TRUE)

## [1] 0.04425837
```

Covariate effects

We also have data on smoking behavior:

- `hsmoke` = indicator of above-average smoking behavior.

```
mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04477612

mean( Y[ hsmoke==1, hsmoke==0] ,na.rm=TRUE)

## [1] 0.03062609

mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

## [1] 0.04400078

mean( Y[ hsmoke==0, hsmoke==0] ,na.rm=TRUE)

## [1] 0.04425837
```

Summarizing densities of subgraphs

There are a lot of probabilities here (four for each covariate)

	xj=0	xj=1
xi=0	0.039	0.039
xi=1	0.038	0.047

Table : gpa

	xj=0	xj=1
xi=0	0.044	0.044
xi=1	0.031	0.045

Table : smoking

Note: Such tables correspond to very rudimentary “blockmodels”:

- an observed categorical covariate divides nodes into “blocks”;
- probability of tie between nodes determined by rates between their blocks.

Interpreting probabilities/rates:
How do rates correspond to nodal preferences?

Summarizing densities of subgraphs

There are a lot of probabilities here (four for each covariate)

	xj=0	xj=1
xi=0	0.039	0.039
xi=1	0.038	0.047

Table : gpa

	xj=0	xj=1
xi=0	0.044	0.044
xi=1	0.031	0.045

Table : smoking

Note: Such tables correspond to very rudimentary "blockmodels":

- an observed categorical covariate divides nodes into "blocks";
- probability of tie between nodes determined by rates between their blocks.

Interpreting probabilities/rates:
How do rates correspond to nodal preferences?

Summarizing densities of subgraphs

There are a lot of probabilities here (four for each covariate)

	xj=0	xj=1
xi=0	0.039	0.039
xi=1	0.038	0.047

Table : gpa

	xj=0	xj=1
xi=0	0.044	0.044
xi=1	0.031	0.045

Table : smoking

Note: Such tables correspond to very rudimentary "blockmodels":

- an observed categorical covariate divides nodes into "blocks";
- probability of tie between nodes determined by rates between their blocks.

Interpreting probabilities/rates:
How do rates correspond to nodal preferences?

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Odds ratios

Odds: Let $\Pr(E)$ be the probability of an event. The odds of E are

$$\text{odds}(E) = \frac{\Pr(E)}{1 - \Pr(E)}$$

Probabilities are between 0 and 1, odds are between 0 and ∞.

The "effect" of a variable on a probability is often described via the odds ratio.

Odds ratio: Let

- $\Pr(E|A)$ = the probability of some event E under condition A
- $\Pr(E|B)$ = the probability of some event E under condition B

The odds ratio is

$$\text{odds}(E : A, B) = \frac{\Pr(E|A)}{1 - \Pr(E|A)} \frac{1 - \Pr(E|B)}{\Pr(E|B)}$$

Note that

$$\text{odds}(E : A, B) = 1 \Rightarrow \Pr(E|A) = \Pr(E|B)$$

Effect of a covariate on a tie

Let $x_i \in \{0, 1\}$ for $i = 1, \ldots, n$ be a binary variable.

- x_i = indicator of high gpa, or
- x_i = smoking status, or
- x_i = indicator of membership to some group.

Let $\Pr(y_{i,j} = 1 | x_i, x_j) = p_{x_i x_j}$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Given a network, we might want to describe the "effect" of x_i and x_j on $y_{i,j}$:

$$\text{odds}(y_{i,j} = 1 | \{x_i = 1, x_j = 1\}, \{x_i = 0, x_j = 1\}) = \frac{p_{11}}{1 - p_{11}} \frac{1 - p_{01}}{p_{01}}$$

Effect of a covariate on a tie

Let $x_i \in \{0, 1\}$ for $i = 1, \dots, n$ be a binary variable.

- $x_i =$ indicator of high gpa, or
- $x_i =$ smoking status, or
- $x_i =$ indicator of membership to some group.

Let $\Pr(y_{i,j} = 1 | x_i, x_j) = p_{x_i x_j}$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Given a network, we might want to describe the "effect" of x_i and x_j on $y_{i,j}$:

$$\text{odds}(y_{i,j} = 1 | \{x_i = 1, x_j = 1\}, \{x_i = 0, x_j = 1\}) = \frac{p_{11}}{1 - p_{11}} \frac{1 - p_{01}}{p_{01}}$$

Effect of a covariate on a tie

Let $x_i \in \{0, 1\}$ for $i = 1, \ldots, n$ be a binary variable.

- $x_i =$ indicator of high gpa, or
- $x_i =$ smoking status, or
- $x_i =$ indicator of membership to some group.

Let $\Pr(y_{i,j} = 1 | x_i, x_j) = p_{x_i x_j}$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Given a network, we might want to describe the "effect" of x_i and x_j on $y_{i,j}$:

$$\text{odds}(y_{i,j} = 1 | \{x_i = 1, x_j = 1\}, \{x_i = 0, x_j = 1\}) = \frac{p_{11}}{1 - p_{11}} \frac{1 - p_{01}}{p_{01}}$$

Effect of a covariate on a tie

Let $x_i \in \{0, 1\}$ for $i = 1, \ldots, n$ be a binary variable.

- x_i = indicator of high gpa, or
- x_i = smoking status, or
- x_i = indicator of membership to some group.

Let $\Pr(y_{i,j} = 1 | x_i, x_j) = p_{x_i x_j}$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Given a network, we might want to describe the "effect" of x_i and x_j on $y_{i,j}$:

$$\text{odds}(y_{i,j} = 1 | \{x_i = 1, x_j = 1\}, \{x_i = 0, x_j = 1\}) = \frac{p_{11}}{1 - p_{11}} \frac{1 - p_{01}}{p_{01}}$$

Effect of a covariate on a tie

Let $x_i \in \{0,1\}$ for $i = 1,\ldots,n$ be a binary variable.

- $x_i =$ indicator of high gpa, or
- $x_i =$ smoking status, or
- $x_i =$ indicator of membership to some group.

Let $\Pr(y_{i,j} = 1 | x_i, x_j) = p_{x_i x_j}$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Given a network, we might want to describe the "effect" of x_i and x_j on $y_{i,j}$:

$$\text{odds}(y_{i,j} = 1 | \{x_i = 1, x_j = 1\}, \{x_i = 0, x_j = 1\}) = \frac{p_{11}}{1 - p_{11}} \frac{1 - p_{01}}{p_{01}}$$

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios

```
p11<-mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)
p01<-mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE)

(p11/(1-p11)) / (p01/(1-p01))

## [1] 1.018447
```

This result says that the odds of a tie are 1.02 times higher under the condition $x_i = 1, x_j = 1$ than $x_i = 0, x_j = 1$.

This result seems to suggest that smokers and non-smokers are equally friendly to smokers. However, the result could be due to

- no effect of smoking *or*
- differential rates of ties among smokers and nonsmokers.

Odds ratios for tie preferences

A better question to ask might be:

> *Does a person's characteristic determine the characteristics of whom they choose as friends?*

The probabilities related to this question condition on the existence of a tie:

$$
\begin{aligned}
\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) &= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= p_{11} \frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}
\end{aligned}
$$

This probability can be interpreted as, for example,

> *What is the probability that a friend of a smoker is another smoker?*

Such a probability is more descriptive of tie preferences.

Odds ratios for tie preferences

A better question to ask might be:

Does a person's characteristic determine the characteristics of whom they choose as friends?

The probabilities related to this question condition on the existence of a tie:

$$\begin{aligned}
\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) &= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1)\Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1)\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= p_{11}\frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}
\end{aligned}$$

This probability can be interpreted as, for example,

What is the probability that a friend of a smoker is another smoker?

Such a probability is more descriptive of tie preferences.

Odds ratios for tie preferences

A better question to ask might be:

> *Does a person's characteristic determine the characteristics of whom they choose as friends?*

The probabilities related to this question condition on the existence of a tie:

$$
\begin{aligned}
\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) &= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1)\Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1)\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= p_{11}\frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}
\end{aligned}
$$

This probability can be interpreted as, for example,

> *What is the probability that a friend of a smoker is another smoker?*

Such a probability is more descriptive of tie preferences.

$\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) =$ probability that a friend of a smoker is another smoker

However, this probability will mostly reflect the (typically low) overall tie density.

To assess the "effect" of x_i on choosing another smoker as a friend, we can look at an appropriate odds ratio:

$$\text{odds}(x_j = 1 : \{y_{i,j} = 1, x_i = 1\}, \{y_{i,j} = 1, x_i = 0\}) = \frac{\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1)}{\Pr(x_j = 0 | y_{i,j} = 1, x_i = 1)} \frac{\Pr(x_j = 0 | y_{i,j} = 1, x_i = 0)}{\Pr(x_j = 1 | y_{i,j} = 1, x_i = 0)}$$

Odds ratios for tie preferences

$\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) =$ probability that a friend of a smoker is another smoker

However, this probability will mostly reflect the (typically low) overall tie density.

To assess the "effect" of x_i on choosing another smoker as a friend, we can look at an appropriate odds ratio:

$$\text{odds}(x_j = 1 : \{y_{i,j} = 1, x_i = 1\}, \{y_{i,j} = 1, x_i = 0\}) = \frac{\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1)}{\Pr(x_j = 0 | y_{i,j} = 1, x_i = 1)} \frac{\Pr(x_j = 0 | y_{i,j} = 1, x_i = 0)}{\Pr(x_j = 1 | y_{i,j} = 1, x_i = 0)}$$

Odds ratios for tie preferences

Recall,

$$\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$
$$= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$
$$= p_{11} \frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

Similarly,

$$\Pr(x_j = 0 | y_{i,j} = 1, x_i = 1) = p_{10} \frac{\Pr(x_j = 0)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

and so

$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1) / \Pr(y_{i,j} = 1 | x_i = 1)}{\Pr(x_j = 0) / \Pr(y_{i,j} = 1 | x_i = 1)}$$
$$= \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}.$$

Odds ratios for tie preferences

Recall,

$$\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

$$= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

$$= p_{11} \frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

Similarly,

$$\Pr(x_j = 0 | y_{i,j} = 1, x_i = 1) = p_{10} \frac{\Pr(x_j = 0)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

and so

$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1) / \Pr(y_{i,j} = 1 | x_i = 1)}{\Pr(x_j = 0) / \Pr(y_{i,j} = 1 | x_i = 1)}$$

$$= \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}.$$

Odds ratios for tie preferences

Recall,

$$\begin{aligned}
\Pr(x_j = 1 | y_{i,j} = 1, x_i = 1) &= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1 | x_i = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= \frac{\Pr(y_{i,j} = 1 | x_j = 1, x_i = 1) \Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)} \\
&= p_{11} \frac{\Pr(x_j = 1)}{\Pr(y_{i,j} = 1 | x_i = 1)}
\end{aligned}$$

Similarly,

$$\Pr(x_j = 0 | y_{i,j} = 1, x_i = 1) = p_{10} \frac{\Pr(x_j = 0)}{\Pr(y_{i,j} = 1 | x_i = 1)}$$

and so

$$\begin{aligned}
\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) &= \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1) / \Pr(y_{i,j} = 1 | x_i = 1)}{\Pr(x_j = 0) / \Pr(y_{i,j} = 1 | x_i = 1)} \\
&= \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}.
\end{aligned}$$

Odds ratios for tie preferences

$$p_{x_1 x_2} = \Pr(\text{ tie } | x_1, x_2)$$
$$\approx \text{ density in the } x_1, x_2 \text{ submatrix}$$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$
$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0) = \frac{p_{01}}{p_{00}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$

The odds ratio is therefore

$$\text{odds ratio}(x_j = 1 | \{y_{i,j} = 1, x_i = 1\}\{y_{i,j} = 1, x_i = 0\}) = \frac{p_{11}/p_{10}}{p_{01}/p_{00}} = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

Odds ratios for tie preferences

$$p_{x_1 x_2} = \Pr(\text{ tie } | x_1, x_2)$$
$$\approx \text{ density in the } x_1, x_2 \text{ submatrix}$$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$
$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0) = \frac{p_{01}}{p_{00}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$

The odds ratio is therefore

$$\text{odds ratio}(x_j = 1 | \{y_{i,j} = 1, x_i = 1\}\{y_{i,j} = 1, x_i = 0\}) = \frac{p_{11}/p_{10}}{p_{01}/p_{00}} = \frac{p_{11}p_{00}}{p_{10}p_{01}}$$

Odds ratios for tie preferences

$$p_{x_1 x_2} = \Pr(\text{ tie } | x_1, x_2)$$
$$\approx \text{ density in the } x_1, x_2 \text{ submatrix}$$

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 1) = \frac{p_{11}}{p_{10}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$
$$\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0) = \frac{p_{01}}{p_{00}} \frac{\Pr(x_j = 1)}{\Pr(x_j = 0)}$$

The odds ratio is therefore

$$\text{odds ratio}(x_j = 1 | \{y_{i,j} = 1, x_i = 1\}\{y_{i,j} = 1, x_i = 0\}) = \frac{p_{11}/p_{10}}{p_{01}/p_{00}} = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11}p_{00}}{p_{10}p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11}p_{00}}{p_{10}p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$
to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$
to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

$$\gamma = \frac{p_{11} p_{00}}{p_{10} p_{01}}$$

This ratio represents

the relative preference of egos with $x = 1$ versus $x = 0$

to tie to alters with $x = 1$.

Interestingly, one can show (homework?)

$$\text{odds ratio}(x_i = 1 | \{y_{i,j} = 1, x_j = 1\}\{y_{i,j} = 1, x_j = 0\}) = \frac{p_{11} p_{00}}{p_{10} p_{01}} = \gamma.$$

This ratio represents

the relative attractiveness of alters with $x = 1$ versus $x = 0$

to egos with $x = 1$.

Odds ratios for tie preferences

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Are there interesting/useful ways to represent numbers in the table?

- In SNA, more interested in relative rates than absolute rates.
- Absolute rates are derivable from relative rates and a baseline, and vice-vera:

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, p_{11}/p_{00}\}$$
$$\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Odds ratios for tie preferences

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Are there interesting/useful ways to represent numbers in the table?

- In SNA, more interested in relative rates than absolute rates.
- Absolute rates are derivable from relative rates and a baseline, and vice-vera:

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, p_{11}/p_{00}\}$$
$$\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Odds ratios for tie preferences

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Are there interesting/useful ways to represent numbers in the table?

- In SNA, more interested in relative rates than absolute rates.
- Absolute rates are derivable from relative rates and a baseline, and vice-vera:

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, p_{11}/p_{00}\}$$
$$\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Odds ratios for tie preferences

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Are there interesting/useful ways to represent numbers in the table?

- In SNA, more interested in relative rates than absolute rates.
- Absolute rates are derivable from relative rates and a baseline, and vice-vera:

$$\begin{aligned}\{p_{00}, p_{01}, p_{10}, p_{11}\} &\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, p_{11}/p_{00}\} \\ &\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}\end{aligned}$$

Odds ratios for tie preferences

	xj=0	xj=1
xi=0	p_{00}	p_{01}
xi=1	p_{10}	p_{11}

Are there interesting/useful ways to represent numbers in the table?

- In SNA, more interested in relative rates than absolute rates.
- Absolute rates are derivable from relative rates and a baseline, and vice-vera:

$$\begin{aligned}\{p_{00}, p_{01}, p_{10}, p_{11}\} &\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, p_{11}/p_{00}\} \\ &\sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}\end{aligned}$$

Interpreting probability ratios

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Basline: p_{00} represents a baseline rate

Relative rates: $p_{01}/p_{00}, p_{10}/p_{00}$ represent relative rates

- p_{01}/p_{00} = density of $0 \to 1$ relative to $0 \to 0$ ("attractiveness of 1's")
- p_{10}/p_{00} = density of $1 \to 0$ relative to $0 \to 0$ ("sociability of 1's")

Odds ratio: $(p_{11}p_{00})/(p_{01}p_{10}) = \gamma$ represents preferences for homophily.

Interpreting probability ratios

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Basline: p_{00} represents a baseline rate

Relative rates: p_{01}/p_{00}, p_{10}/p_{00} represent relative rates

- p_{01}/p_{00} = density of $0 \rightarrow 1$ relative to $0 \rightarrow 0$ ("attractiveness of 1's")
- p_{10}/p_{00} = density of $1 \rightarrow 0$ relative to $0 \rightarrow 0$ ("sociability of 1's")

Odds ratio: $(p_{11}p_{00})/(p_{01}p_{10}) = \gamma$ represents preferences for homophily.

Interpreting probability ratios

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Basline: p_{00} represents a baseline rate

Relative rates: $p_{01}/p_{00}, p_{10}/p_{00}$ represent relative rates

- p_{01}/p_{00} = density of $0 \to 1$ relative to $0 \to 0$ ("attractiveness of 1's")
- p_{10}/p_{00} = density of $1 \to 0$ relative to $0 \to 0$ ("sociability of 1's")

Odds ratio: $(p_{11}p_{00})/(p_{01}p_{10}) = \gamma$ represents preferences for homophily.

Interpreting probability ratios

$$\{p_{00}, p_{01}, p_{10}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{10}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$$

Basline: p_{00} represents a baseline rate

Relative rates: $p_{01}/p_{00}, p_{10}/p_{00}$ represent relative rates

- p_{01}/p_{00} = density of $0 \rightarrow 1$ relative to $0 \rightarrow 0$ ("attractiveness of 1's")
- p_{10}/p_{00} = density of $1 \rightarrow 0$ relative to $0 \rightarrow 0$ ("sociability of 1's")

Odds ratio: $(p_{11}p_{00})/(p_{01}p_{10}) = \gamma$ represents preferences for homophily.

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Interpreting relative rates

You can show (for example) that

$$\frac{p_{01}}{p_{00}} = \frac{\text{odds}(x_j = 1 | y_{i,j} = 1, x_i = 0)}{\text{odds}(x_j = 1)}$$

While this is a ratio of odds, it is not exactly an odds ratio:

- The conditioning events are not complementary.

The ratio still has a reasonable interpretation:

- The ratio can be interpreted as how much the odds of $x_j = 1$ change if you are told that j has a link from a person i with $x_i = 0$.

Friendship example

```
p.smoke<-c(
mean( Y[ hsmoke==0, hsmoke==0] ,na.rm=TRUE),
mean( Y[ hsmoke==1, hsmoke==0] ,na.rm=TRUE),
mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE),
mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)  )

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]  )

## [1] 1.470585

p.gpa<-c(
mean( Y[ hgpa==0, hgpa==0] ,na.rm=TRUE),
mean( Y[ hgpa==1, hgpa==0] ,na.rm=TRUE),
mean( Y[ hgpa==0, hgpa==1] ,na.rm=TRUE),
mean( Y[ hgpa==1, hgpa==1] ,na.rm=TRUE)  )

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Homophily is positive for both smoking and gpa.

Friendship example

```
p.smoke<-c(
mean( Y[ hsmoke==0, hsmoke==0] ,na.rm=TRUE),
mean( Y[ hsmoke==1, hsmoke==0] ,na.rm=TRUE),
mean( Y[ hsmoke==0, hsmoke==1] ,na.rm=TRUE),
mean( Y[ hsmoke==1, hsmoke==1] ,na.rm=TRUE)  )

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]  )

## [1] 1.470585

p.gpa<-c(
mean( Y[ hgpa==0, hgpa==0] ,na.rm=TRUE),
mean( Y[ hgpa==1, hgpa==0] ,na.rm=TRUE),
mean( Y[ hgpa==0, hgpa==1] ,na.rm=TRUE),
mean( Y[ hgpa==1, hgpa==1] ,na.rm=TRUE)  )

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Homophily is positive for both smoking and gpa.

Friendship example

```
p.smoke[1]

## [1] 0.04425837

p.smoke[2]/p.smoke[1]

## [1] 0.6919841

p.smoke[3]/p.smoke[1]

## [1] 0.9941797

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]  )

## [1] 1.470585
```

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

- The baseline rate is low ($p_{00} = 0.04$)
- The rate of ties from nonsmokers to smokers is about the same as that from nonsmokers to nonsmokers ($p_{01}/p_{00} = .99$).
- The rate of ties from smokers to nonsmokers is much lower than that from nonsmokers to nonsmokers ($p_{10}/p_{00} = .69$).
- There is strong homophily ($\gamma = 1.47$)
  - A tie from a smoker is more likely to be to a smoker than a tie from a nonsmoker is.
  - A tie to a smoker is more likely to be from a smoker than a tie to a nonsmoker is.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Friendship example

```
p.gpa[1]

## [1] 0.03903421

p.gpa[2]/p.gpa[1]

## [1] 0.9638469

p.gpa[3]/p.gpa[1]

## [1] 1.008332

(p.gpa[1]*p.gpa[4])  /  (p.gpa[2]*p.gpa[3]  )

## [1] 1.248783
```

Note: It is possible for $p_{01}/p_{00} = p_{10}/p_{00} = 1$, but γ to be large.

- In this case $\gamma = p_{11}/p_{00}$.
- Deviations from 1 indicate heterogeneity in within-group ties.
- Such deviations indicate within group preferences, or homophily.

Logistic regression

A useful tool for describing effects on a binary variable is **logistic regression**

Given

- a binary outcome variable y
- binary explanatory variables x_1, x_2

A logistic regression model for y in terms of x_1, x_2 is

$$\Pr(y = 1|x_1, x_2) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

Based on this, we see that

$$\Pr(y = 0|x_1, x_2) = \frac{1}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

$$\text{odds}(y = 1|x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Logistic regression

A useful tool for describing effects on a binary variable is **logistic regression**

Given

- a binary outcome variable y
- binary explanatory variables x_1, x_2

A logistic regression model for y in terms of x_1, x_2 is

$$\Pr(y = 1 | x_1, x_2) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2}}$$

Based on this, we see that

$$\Pr(y = 0 | x_1, x_2) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2}}$$

$$\text{odds}(y = 1 | x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Logistic regression

A useful tool for describing effects on a binary variable is **logistic regression**

Given

- a binary outcome variable y
- binary explanatory variables x_1, x_2

A logistic regression model for y in terms of x_1, x_2 is

$$\Pr(y = 1|x_1, x_2) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

Based on this, we see that

$$\Pr(y = 0|x_1, x_2) = \frac{1}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

$$\text{odds}(y = 1|x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Logistic regression

A useful tool for describing effects on a binary variable is **logistic regression**

Given

- a binary outcome variable y
- binary explanatory variables x_1, x_2

A logistic regression model for y in terms of x_1, x_2 is

$$\Pr(y = 1|x_1, x_2) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

Based on this, we see that

$$\Pr(y = 0|x_1, x_2) = \frac{1}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

$$\text{odds}(y = 1|x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Logistic regression

A useful tool for describing effects on a binary variable is **logistic regression**

Given

- a binary outcome variable y
- binary explanatory variables x_1, x_2

A logistic regression model for y in terms of x_1, x_2 is

$$\Pr(y = 1|x_1, x_2) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}}$$

Based on this, we see that

$$\begin{aligned}
\Pr(y = 0|x_1, x_2) &= \frac{1}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2}} \\
\text{odds}(y = 1|x_1, x_2) &= \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2) \\
\log \text{odds}(y = 1|x_1, x_2) &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2
\end{aligned}$$

Log-odds ratios in logistic regression

For example,

$$\begin{aligned}
\text{odds}(y = 1|0,0) &= \exp(\beta_0) \\
\text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\
\text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \\
\text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1
\end{aligned}$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

For example,

$$
\begin{aligned}
\text{odds}(y = 1|0,0) &= \exp(\beta_0) \\
\text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\
\text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \\
\text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1
\end{aligned}
$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

For example,

$$\begin{aligned}
\text{odds}(y = 1|0,0) &= \exp(\beta_0) \\
\text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\
\text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \\
\text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1
\end{aligned}$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

For example,

$$\begin{aligned} \text{odds}(y = 1|0,0) &= \exp(\beta_0) \\ \text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\ \text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \\ \text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1 \end{aligned}$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

For example,

$$\begin{aligned}
\text{odds}(y = 1|0,0) &= \exp(\beta_0) \\
\text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\
\text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \\
\text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1
\end{aligned}$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

For example,

$$\begin{aligned}
\text{odds}(y = 1|0,0) &= \exp(\beta_0) \\
\text{odds}(y = 1|1,0) &= \exp(\beta_0 + \beta_1) \\
\text{odds ratio}(y = 1|(1,0),(0,0)) &= \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp\left(\beta_1\right) \\
\text{log odds ratio}(y = 1|(1,0),(0,0)) &= \beta_1
\end{aligned}$$

In logistic regression

- β_1, the "effect" of x_1, represents the log odds ratio $(y = 1|(1,0),(0,0))$
- β_2, the "effect" of x_2, represents the log odds ratio $(y = 1|(0,1),(0,0))$

What about the interaction?

Log-odds ratios in logistic regression

$$\text{odds}(y = 1|x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\text{odds ratio}(y = 1|(1,1),(0,1)) = \frac{\exp(\beta_0 + \beta_1 + \beta_2 + \beta_{12})}{\exp(\beta_0 + \beta_2)} = \exp(\beta_1 + \beta_{12})$$

$$\text{odds ratio}(y = 1|(1,0),(0,0)) = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1)$$

Therefore

$$\frac{\text{odds ratio}(y = 1|(1,1),(0,1))}{\text{odds ratio}(y = 1|(1,0),(0,0))} = \exp(\beta_{12})$$

$$\log \frac{\text{odds ratio}(y = 1|(1,1),(0,1))}{\text{odds ratio}(y = 1|(1,0),(0,0))} = \beta_{12}$$

Log-odds ratios in logistic regression

$$\text{odds}(y = 1|x_1, x_2) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)$$

$$\text{odds ratio}(y = 1|(1,1),(0,1)) = \frac{\exp(\beta_0 + \beta_1 + \beta_2 + \beta_{12})}{\exp(\beta_0 + \beta_2)} = \exp(\beta_1 + \beta_{12})$$

$$\text{odds ratio}(y = 1|(1,0),(0,0)) = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1)$$

Therefore

$$\frac{\text{odds ratio}(y = 1|(1,1),(0,1))}{\text{odds ratio}(y = 1|(1,0),(0,0))} = \exp(\beta_{12})$$

$$\log \frac{\text{odds ratio}(y = 1|(1,1),(0,1))}{\text{odds ratio}(y = 1|(1,0),(0,0))} = \beta_{12}$$

Log-odds ratios in logistic regression

$$o_{x_1 x_2} = \frac{\Pr(y_{1,2} = 1 | x_1, x_2)}{1 - \Pr(y_{1,2} = 1 | x_1, x_2)} = \frac{p_{x_1 x_2}}{1 - p_{x_1 x_2}}$$

	x2=0	x2=1
x1=0	o_{00}	o_{01}
x1=1	o_{10}	o_{11}

Under the logistic regression model

$$\beta_0 = \log o_{00}$$

$$\beta_1 = \log \frac{o_{10}}{o_{00}}$$

$$\beta_2 = \log \frac{o_{10}}{o_{00}}$$

$$\beta_{12} = \log \frac{o_{11}/o_{01}}{o_{10}/o_{00}} = \log \frac{o_{11} o_{00}}{o_{01} o_{10}}$$

How do $\{\beta_0, \beta_1, \beta_2, \beta_{12}\}$ relate to $\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11} p_{00})/(p_{01} p_{10})\}$?

Log-odds ratios in logistic regression

$$o_{x_1x_2} = \frac{\Pr(y_{1,2} = 1|x_1, x_2)}{1 - \Pr(y_{1,2} = 1|x_1, x_2)} = \frac{p_{x_1x_2}}{1 - p_{x_1x_2}}$$

	x2=0	x2=1
x1=0	o_{00}	o_{01}
x1=1	o_{10}	o_{11}

Under the logistic regression model

$$\begin{aligned}
\beta_0 &= \log o_{00} \\
\beta_1 &= \log \frac{o_{10}}{o_{00}} \\
\beta_2 &= \log \frac{o_{10}}{o_{00}} \\
\beta_{12} &= \log \frac{o_{11}/o_{01}}{o_{10}/o_{00}} = \log \frac{o_{11}o_{00}}{o_{01}o_{10}}
\end{aligned}$$

How do $\{\beta_0, \beta_1, \beta_2, \beta_{12}\}$ relate to $\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\}$?

Friendship example

```
Xr<-matrix(hsmoke,nrow(Y),ncol(Y))
Xc<-t(Xr)

xr<-c(Xr)
xc<-c(Xc)
y<-c(Y)

fit<-glm(y~ xr+ xc + xr*xc, family=binomial)

exp(fit$coef)

## (Intercept)          xr          xc       xr:xc
##  0.04630788  0.68225274  0.99391180  1.49277105
```

Do these numbers look familiar?

```
p.smoke[1]

## [1] 0.04425837

p.smoke[2]/p.smoke[1]

## [1] 0.6919841

p.smoke[3]/p.smoke[1]

## [1] 0.9941797

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]  )

## [1] 1.470585
```

Friendship example

```
Xr<-matrix(hsmoke,nrow(Y),ncol(Y))
Xc<-t(Xr)

xr<-c(Xr)
xc<-c(Xc)
y<-c(Y)

fit<-glm(y~ xr+ xc + xr*xc, family=binomial)

exp(fit$coef)

## (Intercept)          xr          xc       xr:xc
##  0.04630788  0.68225274  0.99391180  1.49277105
```

Do these numbers look familiar?

```
p.smoke[1]

## [1] 0.04425837

p.smoke[2]/p.smoke[1]

## [1] 0.6919841

p.smoke[3]/p.smoke[1]

## [1] 0.9941797

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]  )

## [1] 1.470585
```

Friendship example

```r
Xr<-matrix(hsmoke,nrow(Y),ncol(Y))
Xc<-t(Xr)

xr<-c(Xr)
xc<-c(Xc)
y<-c(Y)

fit<-glm(y~ xr+ xc + xr*xc, family=binomial)

exp(fit$coef)

## (Intercept)          xr          xc       xr:xc
##  0.04630788  0.68225274  0.99391180  1.49277105
```

Do these numbers look familiar?

```r
p.smoke[1]

## [1] 0.04425837

p.smoke[2]/p.smoke[1]

## [1] 0.6919841

p.smoke[3]/p.smoke[1]

## [1] 0.9941797

(p.smoke[1]*p.smoke[4])  /  (p.smoke[2]*p.smoke[3]   )

## [1] 1.470585
```

Comparing summaries

If network density is very low,

- $1 - p_{x_i x_j} \approx 1$
- $o_{x_i x_j} = p_{x_i x_j}/(1 - p_{x_i x_j}) \approx p_{x_i x_j}$

and so

	xi=0	xi=1
xj=0	p_{00}	p_{01}
xj=1	p_{10}	p_{11}

$\approx$

	xi=0	xi=1
xj=0	o_{00}	o_{01}
xj=1	o_{10}	o_{11}

Therefore

$$\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\} \approx \{o_{00}, o_{10}/o_{00}, o_{01}/o_{00}, (o_{11}o_{00})/(o_{01}o_{10})\}$$
$$= \{e^{\beta_0}, e^{\beta_1}, e^{\beta_2}, e^{\beta_{12}}\}$$

Comparing summaries

If network density is very low,

- $1 - p_{x_i x_j} \approx 1$
- $o_{x_i x_j} = p_{x_i x_j}/(1 - p_{x_i x_j}) \approx p_{x_i x_j}$

and so

	xi=0	xi=1
xj=0	p_{00}	p_{01}
xj=1	p_{10}	p_{11}

$\approx$

	xi=0	xi=1
xj=0	o_{00}	o_{01}
xj=1	o_{10}	o_{11}

Therefore

$$\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\} \approx \{o_{00}, o_{10}/o_{00}, o_{01}/o_{00}, (o_{11}o_{00})/(o_{01}o_{10})\}$$
$$= \{e^{\beta_0}, e^{\beta_1}, e^{\beta_2}, e^{\beta_{12}}\}$$

Comparing summaries

If network density is very low,

- $1 - p_{x_i x_j} \approx 1$
- $o_{x_i x_j} = p_{x_i x_j}/(1 - p_{x_i x_j}) \approx p_{x_i x_j}$

and so

	xi=0	xi=1
xj=0	p_{00}	p_{01}
xj=1	p_{10}	p_{11}

$\approx$

	xi=0	xi=1
xj=0	o_{00}	o_{01}
xj=1	o_{10}	o_{11}

Therefore

$$\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\} \approx \{o_{00}, o_{10}/o_{00}, o_{01}/o_{00}, (o_{11}o_{00})/(o_{01}o_{10})\}$$
$$= \{e^{\beta_0}, e^{\beta_1}, e^{\beta_2}, e^{\beta_{12}}\}$$

If network density is very low,

- $1 - p_{x_i x_j} \approx 1$
- $o_{x_i x_j} = p_{x_i x_j}/(1 - p_{x_i x_j}) \approx p_{x_i x_j}$

and so

	xi=0	xi=1
xj=0	p_{00}	p_{01}
xj=1	p_{10}	p_{11}

$\approx$

	xi=0	xi=1
xj=0	o_{00}	o_{01}
xj=1	o_{10}	o_{11}

Therefore

$$\{p_{00}, p_{10}/p_{00}, p_{01}/p_{00}, (p_{11}p_{00})/(p_{01}p_{10})\} \approx \{o_{00}, o_{10}/o_{00}, o_{01}/o_{00}, (o_{11}o_{00})/(o_{01}o_{10})\}$$
$$= \{e^{\beta_0}, e^{\beta_1}, e^{\beta_2}, e^{\beta_{12}}\}$$

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
  - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
  - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
    - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
    - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
  - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
    - the excess within-group density, beyond the effect of one group being more active than another.

Undirected data

Now we have

	xj=0	xj=1
xi=0	p_{00}	$p_{01} = p_{10}$
xi=1	$p_{10} = p_{01}$	p_{11}

Now there are only three (unique) numbers in the table.

We can express these as follows:

$$\{p_{00}, p_{01}, p_{11}\} \sim \{p_{00}, p_{01}/p_{00}, p_{11}p_{00}/p_{01}^2\}$$

The interpretation of these is roughly the same as before:

- p_{00} represents a baseline rate (both x's 0)
- p_{01}/p_{00} represents a relative rate (one x 0 versus both x's 0)
- $p_{11}p_{00}/p_{01}^2$ represents a homophily effect - the preference of like for like.
  - the excess within-group density, beyond the effect of one group being more active than another.

Logistic regression for undirected data

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not “sender” and “receiver” effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the “additive” effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Logistic regression for undirected data

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not "sender" and "receiver" effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1 (x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the "additive" effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Logistic regression for undirected data

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not "sender" and "receiver" effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 (x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the "additive" effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Logistic regression for undirected data

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not "sender" and "receiver" effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1 (x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the "additive" effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Logistic regression for undirected data

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not "sender" and "receiver" effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 (x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the "additive" effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Logistic regression for undirected data

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Here, x_1 and x_2 are not "sender" and "receiver" effects, as there are no senders or receivers.

As there is no way to differentiate the effect of x_1 versus that of x_2, we must have $\beta_1 = \beta_2$, and the model becomes

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1 (x_1 + x_2) + \beta_2 x_1 x_2$$

- β_0 represents a baseline rate;
- β_1 represents the "additive" effect of either $x_1 = 1$ or $x_2 = 1$ on the rate;
- β_{12} represents the additional effect of homophily on the rate.

Interpreting coefficients

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1 | 0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1 | 1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1 | 0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1 | 1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1 | x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1 | 0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1 | 1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\log \text{odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

$$\text{log odds}(y = 1|x_1, x_2) = \beta_0 + \beta_1(x_1 + x_2) + \beta_2 x_1 x_2$$

For example, suppose

- $y_{i,j}$ = indicator of friendship;
- x_i = indicator of friendliness.

Under no homophily, i.e. $\beta_{12} = 0$,

$$\text{log odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\text{log odds}(y = 1|1, 1) = \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- not because of homophily,
- but because both people are friendly, instead of one.

Interpreting coefficients

Under positive homophily, i.e. $\beta_{12} > 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1 + \beta_{12} > \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- both people are friendly,
- additionally, friendly people prefer friendly people.

Interpreting coefficients

Under positive homophily, i.e. $\beta_{12} > 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1 + \beta_{12} > \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- both people are friendly,
- additionally, friendly people prefer friendly people.

Interpreting coefficients

Under positive homophily, i.e. $\beta_{12} > 0$,

$$\log \text{odds}(y = 1|0,1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1,1) = \beta_0 + 2\beta_1 + \beta_{12} > \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- both people are friendly,
- additionally, friendly people prefer friendly people.

Interpreting coefficients

Under positive homophily, i.e. $\beta_{12} > 0$,

$$\log \text{odds}(y = 1|0, 1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1, 1) = \beta_0 + 2\beta_1 + \beta_{12} > \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- both people are friendly,
- additionally, friendly people prefer friendly people.

Interpreting coefficients

Under positive homophily, i.e. $\beta_{12} > 0$,

$$\log \text{odds}(y = 1|0,1) = \beta_0 + \beta_1$$
$$\log \text{odds}(y = 1|1,1) = \beta_0 + 2\beta_1 + \beta_{12} > \beta_0 + 2\beta_1$$

The rate is higher under $(x_i = 1, x_j = 1)$ than $(x_i = 1, x_j = 0)$

- both people are friendly,
- additionally, friendly people prefer friendly people.

Computation in R

```
ys<-c(  1*(Y+t(Y)>0)   )

fit<-glm(ys~ xr+ xc + xr*xc, family=binomial)

exp(fit$coef)

## (Intercept)          xr          xc       xr:xc
##  0.07455013  0.84579038  0.84579038  1.45293402
```

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
    - baseline rate;
    - relative probabilities;
    - homophily.
- These summaries are related to logistic regression coefficients.

Summary

- Effects of binary covariates can be described with submatrix densities.
- Submatrix densities can be reparameterized:
  - baseline rate;
  - relative probabilities;
  - homophily.
- These summaries are related to logistic regression coefficients.