Conditional tests

567 Statistical analysis of social networks

Peter Hoff

Statistics, University of Washington
Evaluating statistical models

\[ H : \{y_{i,j}, i \neq j\} \sim \text{binary}(\theta), \text{ for some } \theta \in [0, 1] \]

We would like to evaluate this model, but we don’t know precisely what to expect from it, as we don’t know which is the correct value of \( \theta \), if \( H \) were to be true.

**Problem:** The **null distribution** of \( Y \) depends on the unknown \( \theta \).

**A solution:**
- Perhaps some aspect of the null distribution doesn’t depend on \( \theta \)
- If so, then it could be used to evaluate the null model (for all \( \theta \in [0, 1] \)).

The tool we need to develop this idea further is **conditional probability**
Let $y_1, y_2, y_3 \sim \text{i.i.d. binary}(\theta)$. 

Suppose we are told that $y_1 + y_2 + y_3 = 2$. 

- What is the probability that $y_1 = 1$? 
- What is the probability that $y_1 = y_2 = 1$?
Intuitive calculation

Consider all possible outcomes, before we are told the sum:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(y_1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(y_2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(y_3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Intuitive calculation

Compute the probabilities of each:

| y₁ | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| y₂ | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| y₃ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| Pr | \((1 - \theta)^3\) | \(\theta(1 - \theta)^2\) | \(\theta(1 - \theta)^2\) | \(\theta^2(1 - \theta)\) | \(\theta(1 - \theta)^2\) | \(\theta^2(1 - \theta)\) | \(\theta^2(1 - \theta)\) | \(\theta^3\) |
Now we are told that $y_1 + y_2 + y_3 = 2$:

<table>
<thead>
<tr>
<th>$y_1$</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_2$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$y_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pr</td>
<td>$(1 - \theta)^3$</td>
<td>$\theta(1 - \theta)^2$</td>
<td>$\theta(1 - \theta)^2$</td>
<td>$\theta^2(1 - \theta)$</td>
<td>$\theta(1 - \theta)^2$</td>
<td>$\theta^2(1 - \theta)$</td>
<td>$\theta^2(1 - \theta)$</td>
<td>$\theta^3$</td>
</tr>
</tbody>
</table>

So it seems that having been told $y_1 + y_2 + y_3 = 2$,

- $(1, 1, 0), (1, 0, 1), (0, 1, 1)$ are the only possibilities;
- each of these was equally probable to begin with;
- they should be equally probable given the information.
Hence,
\[
\begin{align*}
\Pr((y_1, y_2, y_3) &= (1, 1, 0) | y_1 + y_2 + y_3 = 2) \\
\Pr((y_1, y_2, y_3) &= (1, 0, 1) | y_1 + y_2 + y_3 = 2) \\
\Pr((y_1, y_2, y_3) &= (0, 1, 1) | y_1 + y_2 + y_3 = 2) \\
\end{align*}
\]
\[
= 1/3
\]

Let’s answer our conditional probability questions:

Suppose we are told that \(y_1 + y_2 + y_3 = 2\).

\begin{itemize}
  \item What is the probability that \(y_1 = 1\)?
    \begin{itemize}
      \item 2/3
    \end{itemize}
  \item What is the probability that \(y_1 = y_2 = 1\)?
    \begin{itemize}
      \item 1/3
    \end{itemize}
\end{itemize}
Conditional probability

Let $A$ and $B$ be two uncertain events.

$$
Pr(B|A) = \frac{Pr(A \text{ and } B)}{Pr(A)}
$$

**Example:** Consider days with non-rainy mornings:

- $B =$ rainy in the evening;
- $A =$ cloudy in the morning.

<table>
<thead>
<tr>
<th></th>
<th>$B$</th>
<th>$B^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>.4</td>
<td>.2</td>
</tr>
<tr>
<td>$A^c$</td>
<td>.1</td>
<td>.3</td>
</tr>
</tbody>
</table>

$$
Pr(B) = Pr(B \cap A) + Pr(B \cap A^c) = .4 + .1 = .5

Pr(A) = Pr(B \cap A) + Pr(B^c \cap A) = .4 + .2 = .6

Pr(B|A) = Pr(B \cap A) / Pr(A) = .4 / .6 = 2/3
$$
Conditional probability

Let $A$ and $B$ be two uncertain events.

$$Pr(B|A) = \frac{Pr(A \text{ and } B)}{Pr(A)}$$

**Example:** A card deck is shuffled and a single card is dealt.

- $B =$ the card is the 3 of hearts.
- $A =$ the card is red.

$$Pr(B|A) = \frac{Pr(A \text{ and } B)}{Pr(A)}$$

$$= \frac{Pr(\text{the card is the 3 of hearts})}{Pr(\text{the card is red})}$$

$$= \frac{1/52}{1/2} = 1/26$$
Conditional probability

Example: Let \( y_1, y_2, y_3 \sim \text{i.i.d. binary}(\theta) \).

- \( B = \{(y_1, y_2, y_3) = (0, 1, 1)\} \)
- \( A = \{y_1 + y_2 + y_3 = 2\} \)

\[
\Pr(B) = (1 - \theta) \times \theta \times \theta = \theta^2 (1 - \theta) \\
\Pr(A) = \Pr((0, 1, 1)) + \Pr((1, 0, 1)) + \Pr((1, 1, 0)) = 3\theta^2 (1 - \theta)
\]

\[
\Pr(B|A) = \frac{\theta^2 (1 - \theta)}{3\theta^2 (1 - \theta)} = \frac{1}{3}.
\]

Note that this probability doesn’t depend on the value of \( \theta \).
A **conditional probability distribution** is an assignment of conditional probabilities to a partition of the outcome space.

Let $B_1, \ldots, B_K$ be a **partition**, so that

- $\Pr(B_j \text{ and } B_k) = 0$;
- $\Pr(B_1 \text{ or } \cdots \text{ or } B_K) = \Pr(B_1) + \cdots \Pr(B_K) = 1$

A conditional probability distribution over $B_1, \ldots, B_K$ given $A$ is simply the collection $\{\Pr(B_k|A), k = 1, \ldots, K\}$. 
**Example:** Let $y_1, y_2, y_3 \sim \text{i.i.d. binary}(\theta)$.

Conditional on $A = \{y_1 + y_2 + y_3 = 2\}$, you should now be able to show that

- $\Pr((y_1, y_2, y_3) = B|A) = 1/3$ if $B$ is either (0,1,1), (1,0,1) or (1,1,0).
- $\Pr((y_1, y_2, y_3) = B|A) = 0$ otherwise

We say the distribution of $y = (y_1, y_2, y_3)$ given $\sum y_i = 2$ is **uniform**, as it assigns equal probabilities to all possible events under the condition.
Conditioning binary sequences

\[ y_1, \ldots, y_m \sim \text{i.i.d. binary}(\theta) \]

What is the conditional distribution of \( y = (y_1, \ldots, y_m) \) given \( s = \sum y_i \)?

Let \( \tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_m) \) be a binary sequence.

\[
Pr(y = \tilde{y}|s) = \frac{Pr(y = \tilde{y} \text{ and } \sum y_i = s)}{Pr(\sum y_i = s)}
\]

First note that this must equal zero if \( \sum \tilde{y}_i \neq s \).
Conditioning binary sequences

\[ y_1, \ldots, y_m \sim \text{i.i.d. binary}(\theta) \]

Let \( \tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_m) \) be a binary sequence with \( \sum \tilde{y}_i = s \).

\[
\Pr(y = \tilde{y}|s) = \frac{\Pr(y = \tilde{y} \text{ and } \sum y_i = s)}{\Pr(\sum y_i = s)}
\]

\[ = \frac{\Pr(y = \tilde{y})}{\Pr(\sum y_i = s)}
\]

\[ = \frac{\theta^s (1 - \theta)^{m-s}}{\binom{m}{s} \theta^s (1 - \theta)^{m-s}} = \frac{1}{\binom{m}{s}}. \]

Recall, \( \binom{m}{s} \) is the number of sequences that have \( s \) ones.

- the probability doesn't depend on \( \theta \);
- the probability is the same for all sequences \( \tilde{y} \) such that \( \sum \tilde{y}_i = s \);

The distribution of \( y|s \) is called a **conditionally uniform** distribution - each possible sequence gets equal probability.
A simulation of $y|s$ can be generated as follows:

0. Put $s$ ones into a bucket, along with $m-s$ zeros.
1. Randomly select a number from the bucket, assign it to $y_1$, and throw it away.
2. Randomly select a number from the bucket, assign it to $y_2$, and throw it away.
   
   : 
   
   m. Select the last number from the bucket and assign it to $y_m$.

This is called (uniform) **sampling without replacement**.

Under this scheme, we will always have $\sum y_i = s$. 
Let’s return to our SRG model for a sociomatrix:

\[
H : Y = \{y_{i,j}, i \neq j\} \sim \text{binary}(\theta), \text{ for some } \theta \in [0, 1]
\]

Under this model, \((y_{1,2}, \ldots, y_{n-1,n})\) forms an i.i.d. binary sequence.

- The conditional distribution of this sequence given \(\sum y_{i,j} = s\) is simply the conditional uniform distribution.
- Knowing \(\sum y_{i,j}\) is the same as knowing \(\bar{y}\).
- The conditional distribution of \(Y\) given \(s\) is sometimes called the Erdos-Reyni graph \((SRG(n, s))\).
- Conditioning on \(s\) is the same as conditioning on \(\bar{y}\).
Simulating from $Y|s$

```r
rY.s <- function(n, s) {
  Y <- matrix(0, n, n); diag(Y) <- NA
  Y[!is.na(Y)] <- sample(c(rep(1, s), rep(0, n*(n-1)-s )))
  Y
}

rY.s(5, 3)

## [1,]   NA   0   0    1   0
## [2,]    0 NA   0   0   0
## [3,]    0   0 NA   0   0
## [4,]    0   0   0 NA   1
## [5,]    1   0   0   0 NA

rY.s(5, 10)

## [1,]   NA   0   1   0   0
## [2,]    1 NA   0   1   1
## [3,]    0   0 NA   0   0
## [4,]    1   0   1 NA   1
## [5,]    0   1   1   1 NA
```
Conditional tests

Suppose $Y \sim SRG(n, \theta)$ for some $\theta \in [0, 1]$. Then

$Y$ should “look like” another sample from $SRG(n, \theta)$

- (but we can’t generate these).

$Y$ should also “look like” another sample from $SRG(n, s)$, where $s = \sum y_{i,j}$

- (we can generate these).

**Conditional evaluation of the SRG**: Given a test statistic $t$, compare $t_{\text{obs}} = t(Y)$ to $\tilde{t} = t(\tilde{Y})$, where $\tilde{Y} \sim SRG(n, \sum y_{i,j})$. 
Example: Monk friendships

\[
\text{mean}(Y, \text{na.rm}=\text{TRUE})
\]
## [1] 0.2875817

\[Cd(Y)\]
## [1] 0.07352941

\[Cd(t(Y))\]
## [1] 0.4044118

\#
\text{nrow}(Y)
## [1] 18

\text{sum}(Y, \text{na.rm}=\text{TRUE})
## [1] 88
Simulated networks

Ysim<-rY.s( nrow(Y), sum(Y,na.rm=TRUE) )
Example: Monk friendships

```r
CD.H <- NULL
for(s in 1:S)
{
  Ysim <- rY.s( nrow(Y), sum(Y, na.rm = TRUE) )
  CD.H <- rbind(CD.H, c(Cd(Ysim), Cd(t(Ysim))))
}
```
Example: Monk friendships

```r
mean(CD.H[,1] <= Cd(Y))
## [1] 0.0024
mean(CD.H[,2] >= Cd(t(Y)))
## [1] 0.025
```

These can be interpreted as *p*-values, but a better thing to say is

- observed outdegree centralization was below the lower 1-percentile of the null distribution;
- observed indegree centralization was above the upper 3-percentile of the null distribution;

The interpretation is that both statistics show evidence against \( H \).
Formal conditional testing

For any test statistic \( t \), consider the following procedure:

1. observe \( Y \)
2. compute \( p = \Pr(t(\tilde{Y}) > t(Y) | H, \sum \tilde{y}_{i,j} = \sum y_{i,j}) \)
3. reject \( H \) if \( p < \alpha \).

If \( Y \sim SRG(n, \theta) \) for some \( \theta \in [0, 1] \) (i.e. \( H \) is true), then

\[
\Pr(\text{reject } H) = \alpha.
\]
Comparison of principled to ad-hoc

CD.H <- NULL
for(s in 1:S)
{
  Ysim <- rY.s( nrow(Y), sum(Y, na.rm=TRUE) )
  # Ysim <- matrix(rbinom(nrow(Y)^2, 1, mean(Y, na.rm=TRUE)), nrow(Y), nrow(Y))
  CD.H <- rbind(CD.H, c(Cd(Ysim), Cd(t(Ysim))))
}

mean(CD.H[,1] <= Cd(Y))
## [1] 0.0022

mean(CD.H[,2] >= Cd(t(Y)))
## [1] 0.0322
Comparison of principled to ad-hoc

```r
CD.H <- NULL
for(s in 1:S)
{
  #Ysim <- rY.s(nrow(Y), sum(Y, na.rm=TRUE))
  Ysim <- matrix(rbinom(nrow(Y)^2, 1, mean(Y, na.rm=TRUE)), nrow(Y), nrow(Y))
  CD.H <- rbind(CD.H, c(Cd(Ysim), Cd(t(Ysim))))
}
mean(CD.H[,1] <= Cd(Y))
## [1] 6e-04
mean(CD.H[,2] >= Cd(t(Y)))
## [1] 0.0214
```
Example: High school friendships

CD.H<-NULL
for(s in 1:S)
{
    Ysim<-rY.s( nrow(Y), sum(Y,na.rm=TRUE) )
    CD.H<-rbind(CD.H, c(Cd(Ysim),Cd(t(Ysim)))))
}
Example: High school friendships

```r
CD.H<-NULL
for(s in 1:S)
{
  # Ysim<-rY.s( nrow(Y), sum(Y,na.rm=TRUE) )
  Ysim<-matrix(rbinom(nrow(Y)^2,1,mean(Y,na.rm=TRUE)),nrow(Y),nrow(Y))
  CD.H<-rbind(CD.H, c(Cd(Ysim),Cd(t(Ysim))))
}
```
Example: High school friendships

```
Sd<-function(Y) { sd(apply(Y,1,sum,na.rm=TRUE)) }
SD.H<-NULL
for(s in 1:S)
{
  Ysim<-rY.s( nrow(Y), sum(Y,na.rm=TRUE) )
  SD.H<-rbind(SD.H, c(Sd(Ysim),Sd(t(Ysim))))
}
```
Example: High school friendships

```r
SD.H <- NULL
for(s in 1:S)
{
  # Ysim <- rY.s( nrow(Y), sum(Y, na.rm=TRUE) )
  Ysim <- matrix(rbinom(nrow(Y)^2, 1, mean(Y, na.rm=TRUE)), nrow(Y), nrow(Y))
  SD.H <- rbind(SD.H, c(Sd(Ysim), Sd(t(Ysim))))
}
```
SRG($n, \theta$) for undirected graphs

```r
n<-10 ; theta<-.1
Y<-matrix(0,n,n)
Y[upper.tri(Y)]<-rbinom( n*(n-1)/2 , 1 , theta )
Y<-Y+t(Y)
diag(Y)<-NA

Y

## [1,] NA  0  0  0  0  0  0  0  0  0
## [2,] 0  NA  0  0  0  0  0  0  0  0
## [3,] 0  0  NA  1  1  1  0  0  0  0
## [4,] 0  0  0  NA  0  0  0  0  0  0
## [5,] 0  0  1  0  NA  0  0  1  0  0
## [6,] 0  0  1  0  0  NA  0  0  0  0
## [7,] 0  0  1  0  0  0  NA  0  0  0
## [8,] 0  0  0  0  1  0  0  NA  0  0
## [9,] 0  0  0  0  0  0  0  NA  0  0
## [10,] 0  0  0  0  0  0  0  0  NA  NA

sum(Y)

## [1] NA
```

sum(Y)

## [1] NA
SRG($n, s$) for undirected graphs

```r
n<-10 ; s<-10
Y<-matrix(0,n,n)
Y[upper.tri(Y)]<-sample( c(rep(1,s),rep(0,n*(n-1)/2 - s)) )
Y<-Y+t(Y)
diag(Y)<-NA
Y
```

```

[1,]  NA  0  0  0  1  0  1  1  1  0
[2,]  0  NA  0  0  0  0  0  0  0  0
[3,]  0  0  NA  1  1  1  0  0  0  0
[4,]  0  0  1  NA  0  0  0  0  0  0
[5,]  1  0  1  0  NA  0  0  1  0  0
[6,]  0  0  1  0  0  NA  0  0  0  0
[7,]  1  0  0  0  0  0  NA  0  0  0
[8,]  1  0  0  0  1  0  0  NA  0  1
[9,]  1  0  0  0  0  0  0  NA  1
[10,] 0  0  0  0  0  0  0  1  1  NA

sum(Y)
```

```
##[1] NA
```

```r
```
Permutation tests for covariate effects

Recall the high-school girls friendship data:

- \( y_{i,j} \) = indicator of friendship, among 144 students;
- \( x_i \) = indicator of above-average smoking behavior.

<table>
<thead>
<tr>
<th></th>
<th>( x_j = 0 )</th>
<th>( x_j = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_i = 0 )</td>
<td>0.044</td>
<td>0.044</td>
</tr>
<tr>
<td>( x_i = 1 )</td>
<td>0.031</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Table: friendship rates between smoking categories

```r
Xrow<-outer(x,rep(1,n))
Xcol<-outer(rep(1,n),x)
Xint<-outer(x,x)

fit<-glm( c(Y) ~ c(Xrow) + c(Xcol) + c(Xint) ,family=binomial)
fit$coef
# (Intercept) c(Xrow) c(Xcol) c(Xint)
# -3.072443033 -0.382355105 -0.006106814 0.400634157

exp(fit$coef)
# (Intercept) c(Xrow) c(Xcol) c(Xint)
# 0.04630788 0.68225274 0.99391180 1.49277105
```
SRG null distribution

\[
\text{ebeta.obs} \leftarrow \text{exp} \left( \text{fit$coef} \right)
\]

\[
\text{ebeta.obs}
\]

## (Intercept)  c(Xrow)  c(Xcol)  c(Xint)
## 0.04630788  0.68225274  0.99391180  1.49277105

\[
\text{EBETA.sim} \leftarrow \text{NULL}
\]

\[
\text{for}(s \in 1:1000)
\]

\[
\{
\text{Ysim} \leftarrow \text{rY.s( nrow(Y), sum(Y,na.rm=TRUE) )}
\text{beta.sim} \leftarrow \text{glm( c(Ysim) ~ c(Xrow) + c(Xcol) + c(Xint), family=binomial)$coef}
\text{EBETA.sim} \leftarrow \text{rbind(EBETA.sim,exp(beta.sim) )}
\}
\]

\[
\text{head(EBETA.sim)}
\]

## (Intercept)  c(Xrow)  c(Xcol)  c(Xint)
## [1,] 0.04649499  0.8854832  0.9307677  1.0051026
## [2,] 0.04295135  1.0370672  0.9243420  1.0544853
## [3,] 0.04054054  1.0466761  1.1824768  0.7870772
## [4,] 0.04500000  0.9102323  0.9429514  1.0795469
## [5,] 0.04444048  1.0213677  0.8745243  1.0442391
SRG null comparisons

The null distribution we are using is a **conditional** null distribution:

- It is conditional on the observed number of ties $\sum_i i \neq j y_{i,j}$;
- It is *also* conditional on the observed values of $x$, smoking behavior.

This distribution tests the following *joint* model for relational and nodal data:

**Model:** $\{Y, x\} \sim p(Y, x)$

**Null Hypothesis:**

- $p(Y, x) = p(Y) \times p(x)$, (ties are independent of covariates) **and**
- $p(Y)$ is a SRG($\theta$) distribution, for some $\theta$.

**Null Distribution:** Test statistics are generated by

- simulating $\tilde{Y}$ from the SRG conditional on $\sum \tilde{y}_{i,j} = \sum y_{i,j}$;
- simulating $\tilde{x}$ conditional on $\tilde{x} = x$. 
SRG null comparisons

...
As compared to any SRG distribution for $Y$ independent of $x$,

- the data show lower smoker sociability ($\beta_r$)
- the data show higher homophily ($\beta_{int}$).

We reject the following hypothesis:

**Null Hypothesis:**

- $p(Y, x) = p(Y) \times p(x)$, (ties are independent of covariates) **and**
- $p(Y)$ is a SRG($\theta$) distribution, for some $\theta$.

But what are we rejecting?

- Are we rejecting because $x$ and $Y$ are not independent?
- Are we rejecting the SRG for $Y$?

The rejection of the test isn’t particularly compelling if we already suspect the SRG to be a poor model.
Tests based on exchangeability

Consider instead the following nonparametric null hypothesis:

**Null Hypothesis:**
- \( p(\mathbf{Y}, \mathbf{x}) = p(\mathbf{Y}) \times p(\mathbf{x}) \), (ties are independent of covariates) and
- \( x_1, \ldots, x_n \sim \text{i.i.d. } p(x) \) for some distribution \( p(x) \).

Consider simulating values of \((\tilde{x}, \tilde{Y})\) the null distribution:
- Can’t do it unconditionally, don’t know \( p(\mathbf{Y}) \) or \( p(\mathbf{x}) \).
- What about conditionally?

**Null Distribution:** Condition on \( \tilde{Y} = Y \), \( \text{sort}(\tilde{x}_1, \ldots, \tilde{x}_n) = \text{sort}(x_1, \ldots, x_n) \).
- Simulate \( \tilde{Y} \) conditional on \( \tilde{Y} = Y \);
- Simulate \( \tilde{x} \) by permuting the entries of \( x \).

**Null scenario:** The scenario that is being mimicked here is \( Y \) being fixed, \( x \) determined independent of \( Y \).
Permutation null distribution

```r
EBETA.psim<-NULL
for(s in 1:1000)
{
  xs<-sample(x)
  Xsrow<-outer(xs,rep(1,n)) ; Xscol<-t(Xsrow) ; Xsint<-Xsrow*Xscol

  beta.sim<-glm( c(Y) ~ c(Xsrow) + c(Xscol) + c(Xsint) , family=binomial)$coef
  EBETA.psim<-rbind(EBETA.psim,exp(beta.sim) )
}

head(EBETA.psim)

## (Intercept) c(Xsrow) c(Xscol) c(Xsint)
## [1,] 0.04724409 0.9921875 0.8137748 0.9586459
## [2,] 0.04855761 0.8565281 0.7188310 1.5099572
## [3,] 0.03814086 1.1900643 1.0959716 0.9323231
## [4,] 0.04500000 1.1900643 1.0959716 0.9323231
## [5,] 0.04500000 0.9289176 0.9335937 1.0553645
## [6,] 0.05006280 0.8813055 0.7721326 1.0520387
```
Permutation null comparisons

- Intercept: Values range from 0.03 to 0.06, with a peak around 0.05.
- Sender smoke: Values range from 0.6 to 1.6, with a peak around 1.0.
- Receiver smoke: Values range from 0.5 to 2.0, with a peak around 1.0.
- Interaction: Values range from 0.5 to 2.0, with a peak around 1.0.
SRG null comparisons

intercept
0.030 0.040 0.050 0.060
0 50 100 150
sender smoke
0.6 0.8 1.0 1.2 1.4 1.6
0 1 2 3 4
receiver smoke
0.5 1.0 1.5
0 1 2 3 4
interaction
0.5 1.0 1.5 2.0
0.0 1.0 2.0
Comparison

While the conclusions for this dataset are basically the same, the evidence against the permutation null is generally weaker than against the SRG null.

- The SRG null makes stronger assumptions (that are generally false);
- The permutation null test requires conditioning on much of the data (which lowers power).

Recommendations:
If your goal is just to reject a null, then use the SRG null.

If you’d like to make more meaningful conclusions, use the permutation null.

Limitations: Permutation approaches only test coarse hypotheses:
- can test for no effect of a nodal covariate;
- can’t test for homophily, in the presence of row and column effects.