

GORIC Exercise 1: Preliminary

Before you look at the data, specify your hypothesis.

As a reminder, the 5 groups are:

1. disadvantaged inner city,
2. advantaged suburban,
3. advantaged rural,
4. disadvantaged rural,
5. disadvantage Spanish speaking.

Let us say, our hypotheses is that:

- ▶ In rural areas (groups 3 and 4), the difference in knowledge of numbers is lower than in other regions.
- ▶ Advantaged regions (groups 2 and 3) improve more than disadvantaged regions.

GORIC Exercise 1: A-prior hypothesis

Let us say, our hypotheses is that:

- ▶ In rural areas (groups 3 and 4), the difference in knowledge of numbers is lower than in other regions.

$$\{\mu_3, \mu_4\} < \{\mu_1, \mu_2, \mu_5\}.$$

Thus, among other things $\mu_3 < \mu_2$ (i.e., $\mu_2 > \mu_3$).

- ▶ Advantaged regions (groups 2 and 3) improve more than disadvantaged regions.

$$\{\mu_2, \mu_3\} > \{\mu_1, \mu_4, \mu_5\}.$$

Thus, among other things $\mu_3 > \mu_4, \mu_2 > \mu_1, \mu_2 > \mu_5$.

Hence, a possible hypothesis could be:

$$H1 : \mu_3 > \mu_4, \mu_2 > \mu_1, \mu_2 > \mu_5, \mu_2 > \mu_3$$

GORIC Exercise 1: Data

Before you start the analysis, do checks on your data.

For example,

- ▶ make sure that factors are coded as factors.
- ▶ make sure that missing data are coded as missings.

GORIC Exercise 1: Null hypothesis test

```
# First, we need the R object with unconstrained estimates  
lm_fit_sesam <- lm((postnumb-prenumb) ~ site-1, data = data)
```

```
# p-value null hypothesis test(s)  
summary(lm_fit_sesam)
```

- ▶ Each mean is significantly different from 0.
- ▶ Hypothesis that all means are zero is rejected as well:
F-statistic: 46.08 on 5 and 235 DF, p-value: < 2.2e-16.
- ▶ But: We do not know anything yet about the orderings expected on
forehand as specified in our theories/expectations/hypotheses, here:
 $H1: \mu_3 > \mu_4, \mu_2 > \mu_1, \mu_2 > \mu_5, \mu_2 > \mu_3$

GORIC Exercise 1: GORIC

The hypothesis of interest, $H1 : \mu_3 > \mu_4, \mu_2 > \mu_1, \mu_2 > \mu_5, \mu_2 > \mu_3$, is reflected by (using the name and levels of the factor):

```
H1_sesam <- 'site3 > site4; site2 > site1; site2 > site5; site2
```

The GORIC weights are obtained via:

```
set.seed(123) # Set seed value
output <- goric(lm_fit_sesam, H1_sesam,
                comparison = 'complement')
summary(output)
```

- ▶ The order-restricted hypothesis '*H1_sesam*' has 5.477 times more support than its complement.
- ▶ We can now quantify the support for our hypothesis/-es of interest.