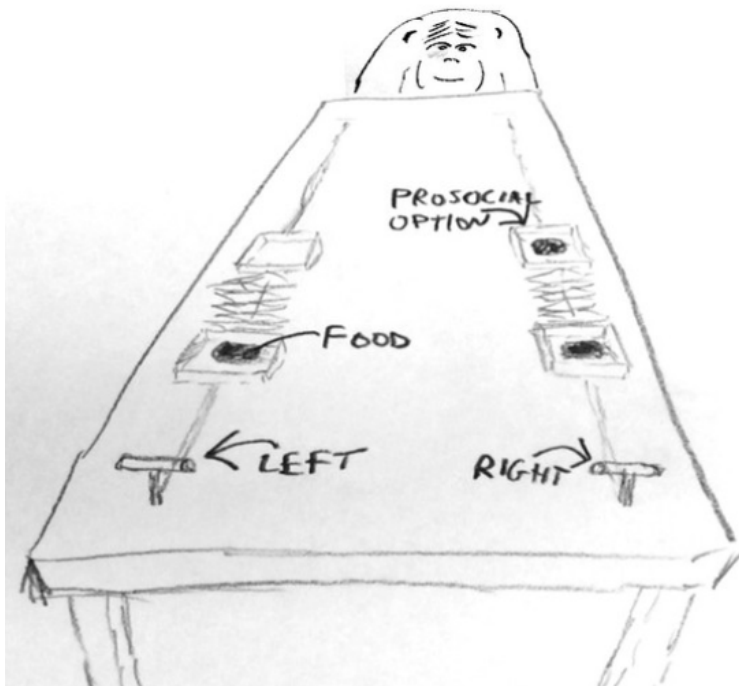# PRESENTATION CONTENTS

1__ Case Study 1 : Is Chimpanzee Sociable?

2__ Case study 2 : Marriage age, rate and Divorce rate

3__ Case study 3 : Oceanic Tool Complexity

Stan

StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 11

1. Experiment
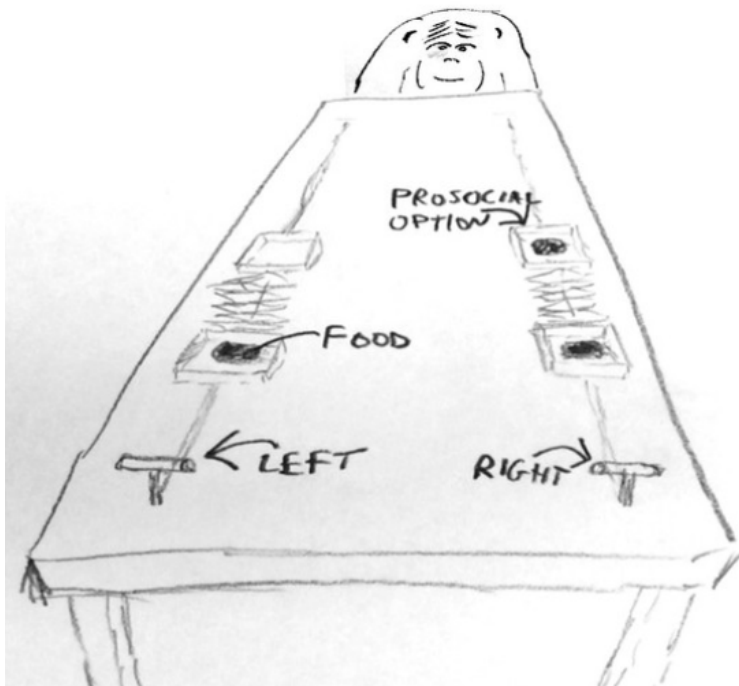
1) As in the figure, there is a chimpanzee(chimp A) which has various options.

2) In the left side there is a food only in chimp A's side, while in the right both side of plates have food.

3) There will be two conditions and two options: whether there is a partner(chimp B) or not, and whether the prosocial option(chimp A and B all can eat the food) is in the right or left.

4) In the given conditions, chimp A should pull one of the lever, and the proportion of pulling left lever will be recorded as the data.

2. Purpose

We want to see the interaction of one's handedness and the condition to the outcome(which side of the lever chimp A select).

Case Study 1: Is Chimpanzee Sociable?

Ch. 11

3. Method

1) There are four possible treatments as the table below:

	Prosocial in	
	Right (R)	**Left (L)**
Partner O (P)	R/P	L/P
Partner X (N)	R/N	L/N

2) We will use *logit* link to calculate the proportion.

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \beta_{\text{TREATMENT}[i]}$$
$$\alpha_j \sim \text{to be determined}$$
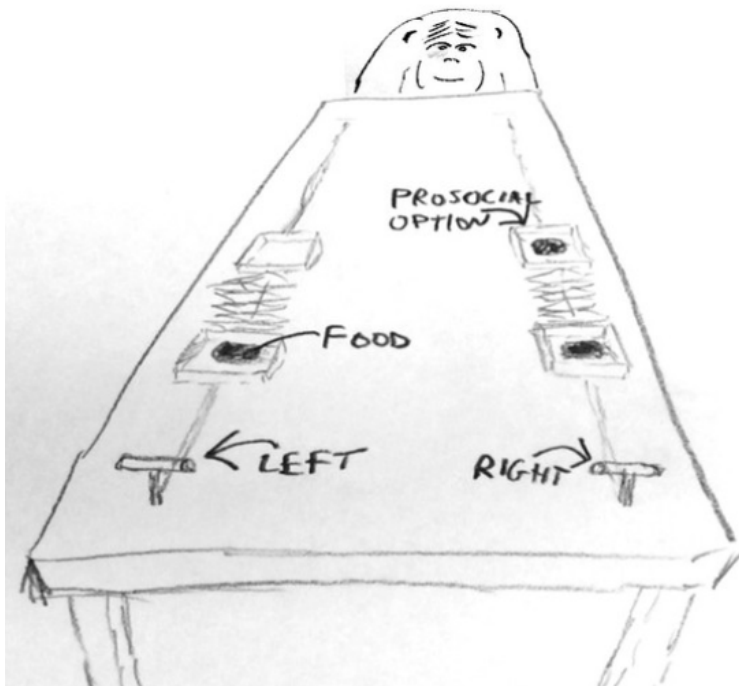$$\beta_k \sim \text{to be determined}$$
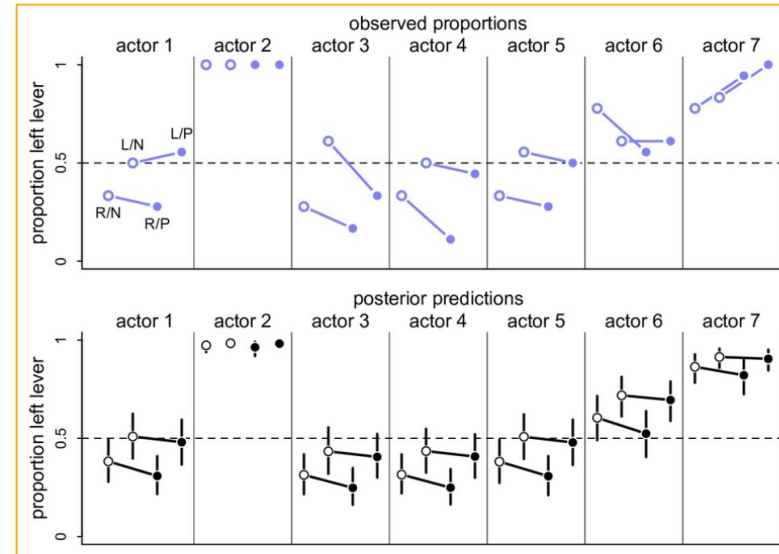
m11.4

α : Handedness (measurement error)
β : condition of partner

StanKorea

# Case Study 1: Is Chimpanzee Sociable?

Ch. 11

4. Result

Raw data

Posterior distribution

1) Actor 2 shows always left selection : No worth

2) Actor 1 shows pretty prosocial tendency, but chimps do not respond always in the same way. That doesn't mean that chimps are bull shit asocial assholes but **their selections are not much affected by partners but by their handedness**.

3) In the posterior distribution, such tendency becomes stronger that the conditions are not that much effective; just their handedness control them.

StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 11
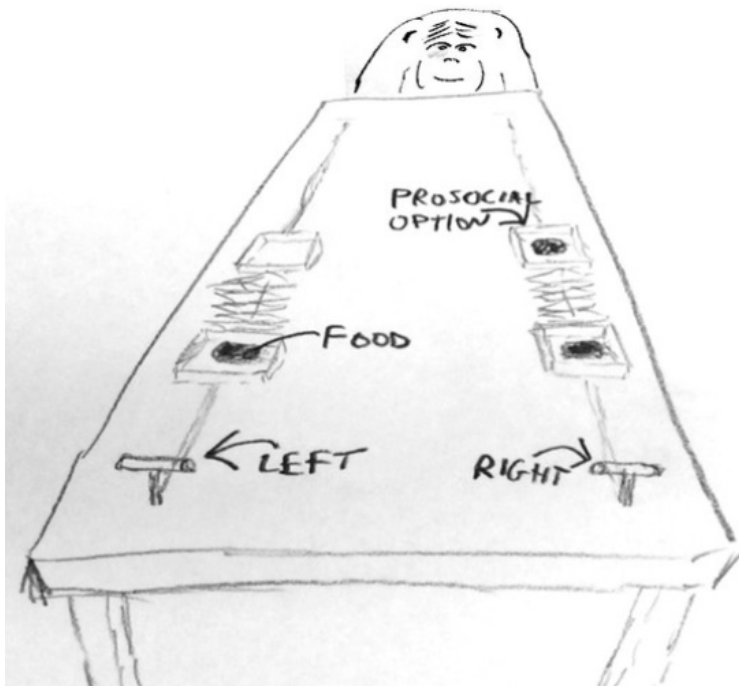
4. Result

4) To make the point clearer, the experiment result was compared with no-interaction of handedness and treatments model.

5) As seen in the dLOO value, which corresponds to WAIC, **there's almost NO interaction between handedness and treatments**.

```
m11.5 <- ulam(
    alist(
        pulled_left ~ dbinom( 1 , p ) ,
        logit(p) <- a[actor] + bs[side] + bc[cond] ,
        a[actor] ~ dnorm( 0 , 1.5 ),
        bs[side] ~ dnorm( 0 , 0.5 ),
        bc[cond] ~ dnorm( 0 , 0.5 )
    ) ,
    data=dat_list2 , chains=4 , log_lik=TRUE )
```

All the parameters are separated with no interaction.

R code 11.19
```
compare( m11.5 , m11.4 , func=LOO )
```

	LOO	pLOO	dLOO	weight	SE	dSE
m11.5	531.2	7.9	0.0	0.66	19.17	NA
m11.4	532.6	8.7	1.4	0.34	19.01	1.28

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \beta_{\text{TREATMENT}[i]}$$
$$\alpha_j \sim \text{to be determined}$$
$$\beta_k \sim \text{to be determined}$$

m11.4

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \beta_{\text{SIDE}[i]} + \beta_{\text{COND}[i]}$$
$$\alpha_j \sim \text{to be determined}$$
$$\beta_k \sim \text{to be determined}$$

m11.5

StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 13

$$L_i \sim \text{Binomial}(1, p_i)$$

$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \gamma_{\text{BLOCK}[i]} + \beta_{\text{TREATMENT}[i]}$$

$$\beta_j \sim \text{Normal}(0, 0.5) \quad , \text{for } j = 1..4$$

$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma_\alpha) \quad , \text{for } j = 1..7$$

$$\gamma_j \sim \text{Normal}(0, \sigma_\gamma) \quad , \text{for } j = 1..6$$

$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$

$$\sigma_\alpha \sim \text{Exponential}(1)$$
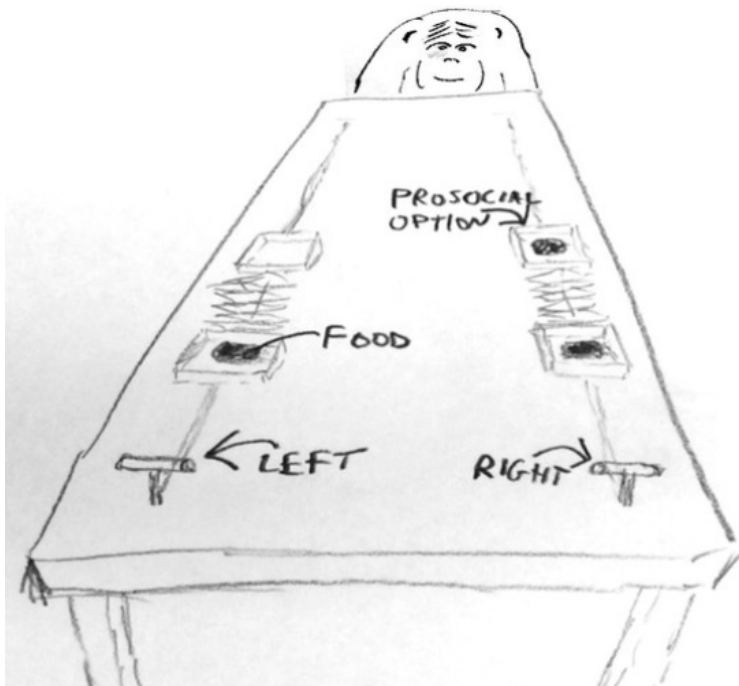
$$\sigma_\gamma \sim \text{Exponential}(1)$$

varying intercepts on block

m13.6

5. New method

1) Experiments are done for six days.

2) We add *varying intercepts* term to our model.

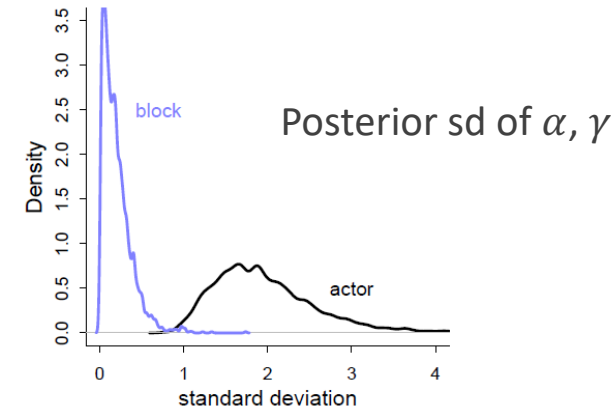# Case Study 1: Is Chimpanzee Sociable?
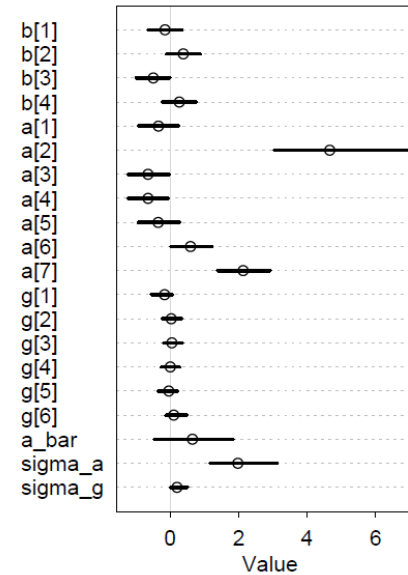
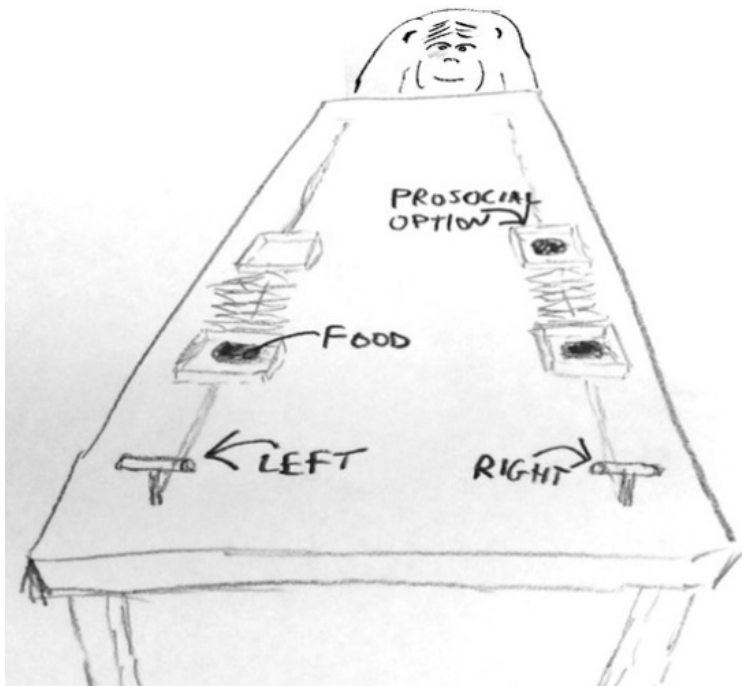Ch. 13

6. New result

Posterior mean & interval

Posterior sd of α, γ

1) Estimated variation among actors is a lot larger than the estimated variation among blocks.

2) This means **the chimpanzees vary, but the blocks are all the same**.

StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 13

6. New result

4) As seen in pWAIC value, we added 7 more parameters but only about 2.5 more parameters are effective.

5) Incorporating block intercept terms has no benefits.

```
set.seed(14)
m13.5 <- ulam(
    alist(
        pulled_left ~ dbinom( 1 , p ) ,
        logit(p) <- a[actor] + b[treatment] ,
        b[treatment] ~ dnorm( 0 , 0.5 ),
        a[actor] ~ dnorm( a_bar , sigma_a ),
        a_bar ~ dnorm( 0 , 1.5 ),
        sigma_a ~ dexp(1)
    ) , data=dat_list , chains=4 , cores=4 , log_lik=TRUE )
```

R code
13.24

```
compare( m13.4 , m13.5 )
```

	WAIC	pWAIC	dWAIC	weight	SE	dSE
m13.5	531.0	8.5	0	0.73	19.23	NA
m13.4	532.9	10.9	2	0.27	19.39	1.64

StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 14

$$L_i \sim \text{Binomial}(1, p_i)$$

$$\text{logit}(p_i) = \gamma_{\text{TID}[i]} + \alpha_{\text{ACTOR}[i],\text{TID}[i]} + \beta_{\text{BLOCK}[i],\text{TID}[i]}$$

Mean treatment effects   *Each actor in each treatment*   *Each block in each treatment*

$$\begin{bmatrix} \alpha_{j,1} \\ \alpha_{j,2} \\ \alpha_{j,3} \\ \alpha_{j,4} \end{bmatrix} \sim \text{MVNormal} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{S}_{\text{ACTOR}} \right)$$

$$\begin{bmatrix} \beta_{j,1} \\ \beta_{j,2} \\ \beta_{j,3} \\ \beta_{j,4} \end{bmatrix} \sim \text{MVNormal} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{S}_{\text{BLOCK}} \right)$$

m14.2

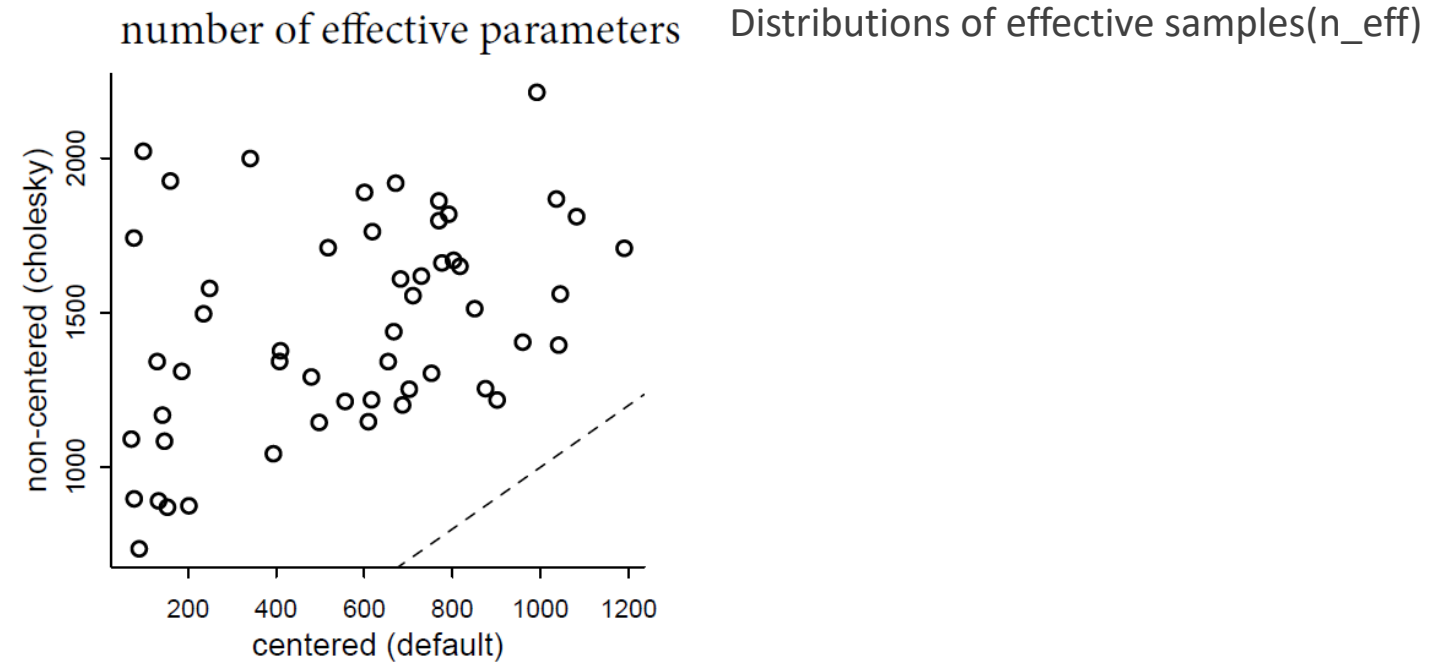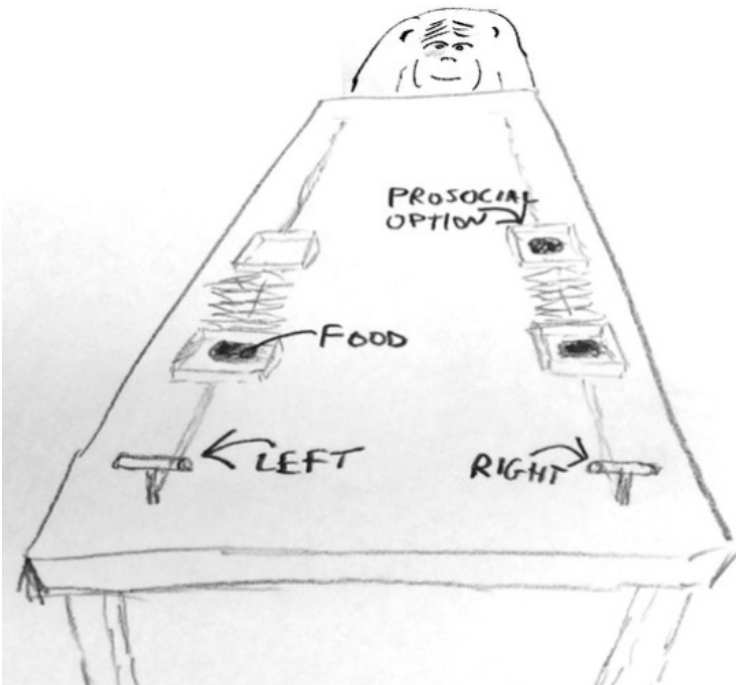7. **Another new method**

   1) **Cross-classified varying slope model**
   – each actor in each treatment, each block in each treatment

   2) We use matrix representation.

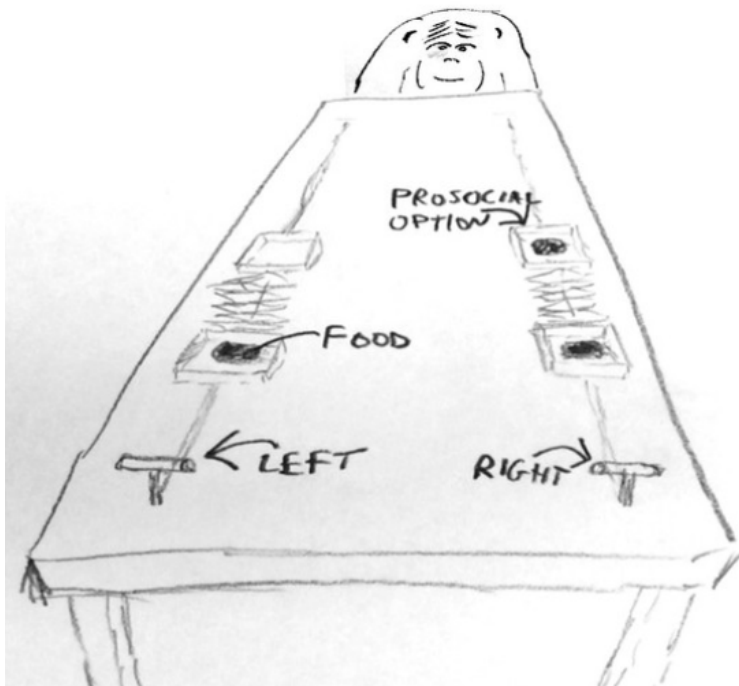StanKorea

Case Study 1: Is Chimpanzee Sociable?

Ch. 14

8. Another new result

number of effective parameters — Distributions of effective samples(n_eff)

1) Both models arrive at equivalent inferences

2) But the non-centered version **samples much more efficiently.**
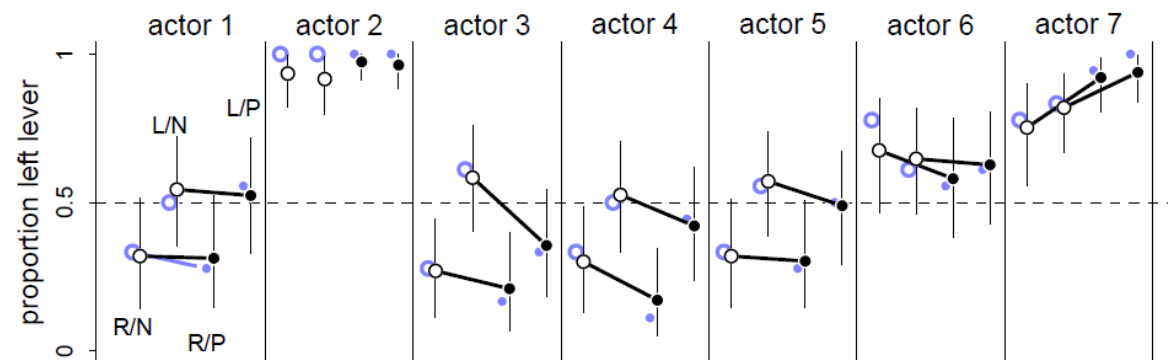
StanKorea

# Case Study 1: Is Chimpanzee Sociable?

Ch. 14

8. Another new result

4) Black lines are posterior predictions and blue lines are the raw data for m14.3.

5) The most obvious difference from earlier is that the model accommodates a lot more variation among individuals.

	mean	sd	5.5%	94.5%	n_eff	Rhat
sigma_actor[1]	1.39	0.49	0.80	2.24	906	1
sigma_actor[2]	0.92	0.38	0.44	1.64	1060	1
sigma_actor[3]	1.86	0.57	1.14	2.89	1191	1
sigma_actor[4]	1.59	0.66	0.86	2.81	1148	1

Case study 2 : Marriage age, rate and Divorce rate

Ch. 5

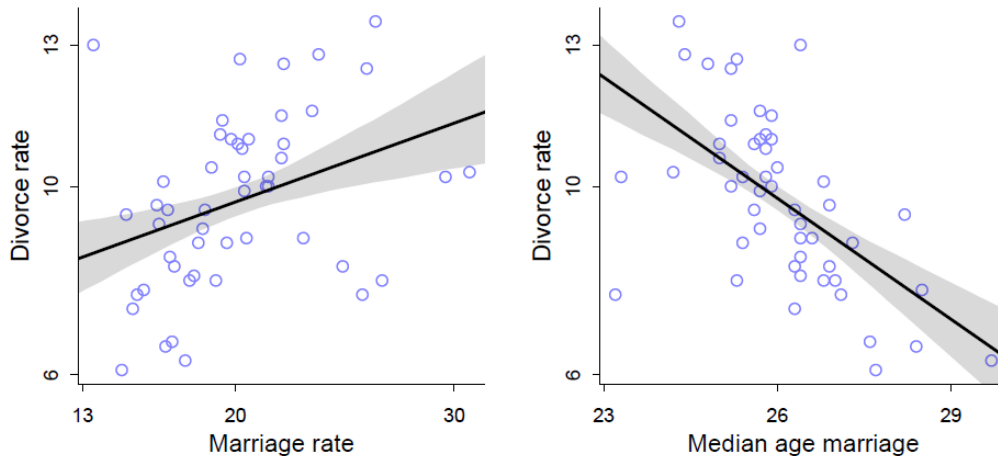Median age of marriage — Marriage rate

A → M

D
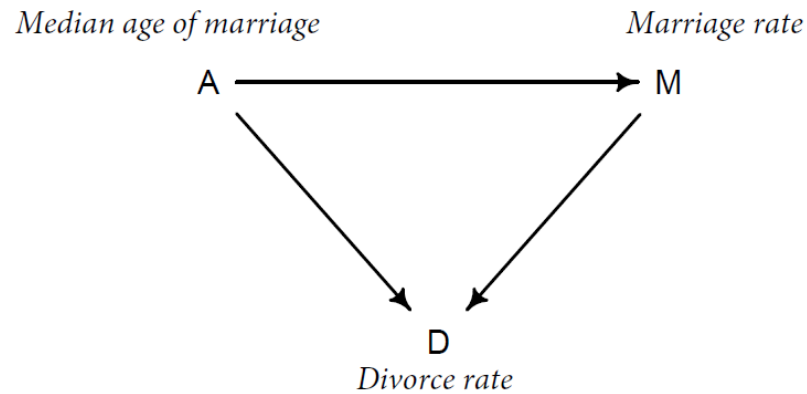
Divorce rate

1. Experiment

As in the figure below, the association between marriage rate (M), divorce rate (D) and median age of marriage (A) are searched. The **multiple regression** will be used to know the genuine correlation and causation between the predictors.
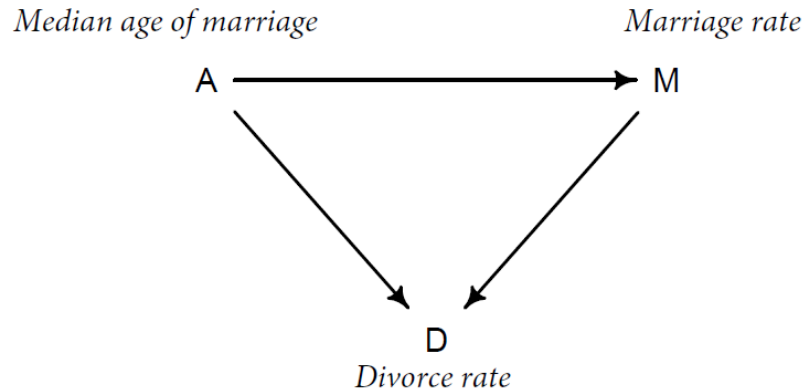
2. Purpose

Although the M, A and D have the association as in the chart, it cannot say the real correlation between them. Spurious correlation can be hidden, hindering the correct causal inference. We will search for

1) the value of a predictor, once we know the other predictors

2) how each predictor is associated with outcome, once we know all the other predictors

StanKorea

Median age of marriage — A ⟶ M — Marriage rate
A, M both point to D — Divorce rate

```
m5.1 <- quap(
   alist(
      D ~ dnorm( mu , sigma ) ,
      mu <- a + bA * A ,
      a ~ dnorm( 0 , 0.2 ) ,
      bA ~ dnorm( 0 , 0.5 ) ,
      sigma ~ dexp( 1 )
   ) , data = d )
m5.3 <- quap(
   alist(
      D ~ dnorm( mu , sigma ) ,
      mu <- a + bM*M + bA*A ,
      a ~ dnorm( 0 , 0.2 ) ,
      bM ~ dnorm( 0 , 0.5 ) ,
      bA ~ dnorm( 0 , 0.5 ) ,
      sigma ~ dexp( 1 )
   ) , data = d )
precis( m5.3 )
```

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_A A_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m5.1

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_M M_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
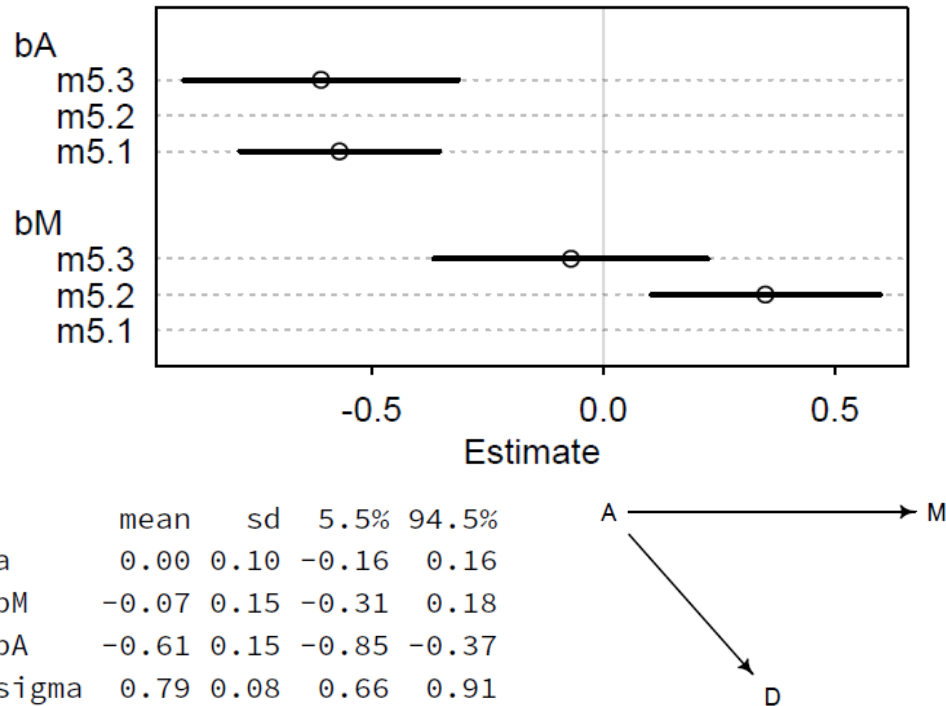$$\sigma \sim \text{Exponential}(1)$$

m5.2

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_M M_i + \beta_A A_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m5.3

3. Method – Multiple regression

1) As seen in the DAG, A, M, D are associated while we cannot know the genuine correlation between them since **whether the path M→D is closed or not does not effect on the result**.

2) For achieving the purpose, we will make three models as in the code; m5.1, 5.2, 5.3 each indicates the code assuming that divorce rate is correlated with median age of marriage only, marriage rate only and both. The last one is **multiple regression**, using two or more variables to predict posterior.

3) The posterior predictions will then be plotted in three other ways : predictor residual plots, counterfactual plots and posterior prediction plots.

StanKorea

Case study 2 : Marriage age, rate and Divorce rate

Ch. 5


```
         mean   sd  5.5% 94.5%
a        0.00 0.10 -0.16  0.16
bM      -0.07 0.15 -0.31  0.18
bA      -0.61 0.15 -0.85 -0.37
sigma    0.79 0.08  0.66  0.91
```

## 4. Result

1) As seen in the chart, the β_A, slope of the marriage age, shows negative posterior distribution both in m5.1 and 5.3 while the β_M, slope of the marriage rate, moves to the zero in multiple regression compared to the single regression.

2) Such shift indicates that **the marriage rate cannot be the direct common cause of the divorce rate**. The DAG should be modified as in the bottom; no direct causal relationship between M and D. They are just associated *through* median marriage age.

3) This is why the multiple regression is useful; **it can show us the genuine causal relationship between predictors and posterior.**

4) Once we know the A, M has little value to produce the posterior distribution. In contrast, A is still valuable in case where we know the M to get D.

StanKorea
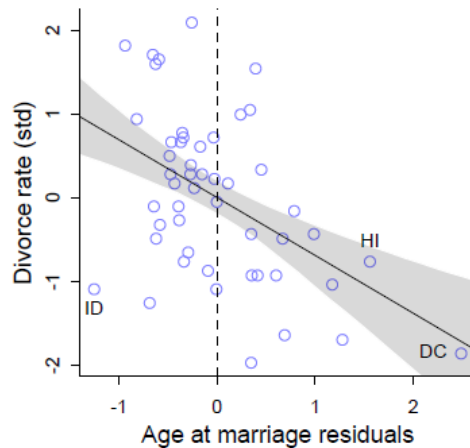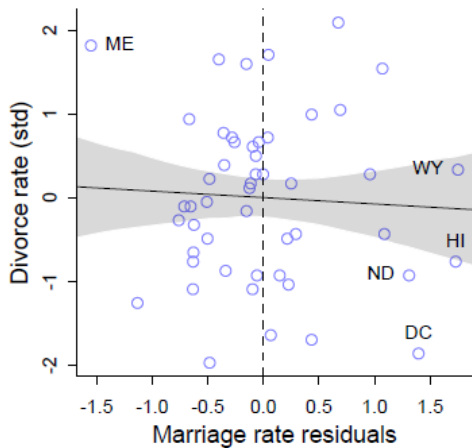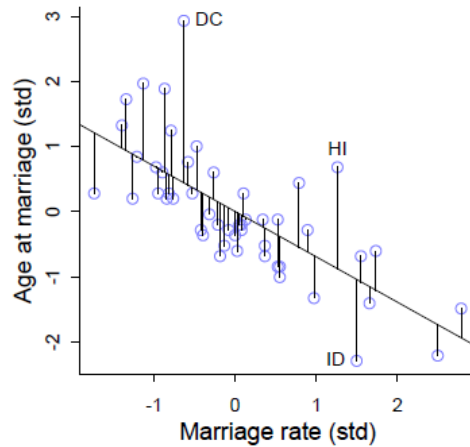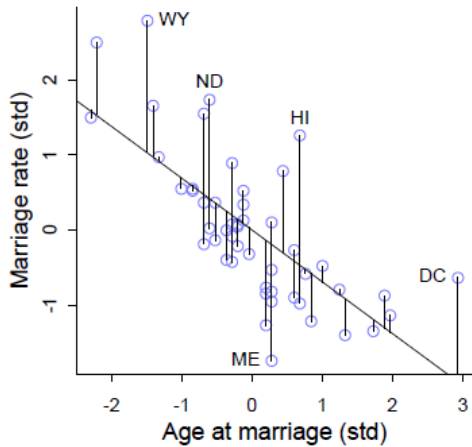
5. Posterior Predictions Plotting

1) We have lots of options to plot the posterior predictions, so we will introduce 3 representative methods : predictor residual, counterfactual and posterior prediction plots.

2) The graph beside shows **predictor residual plots**. Regressing the predictor on other predictors, we can calculate the residual of the predictor and regress outcome on residuals.

3) The regression of the marriage rate residuals to divorce rate shows no association while the age residuals show clear negative relationship with divorce rate.

4) Predictor residual plots are **very intuitive to evaluate the predictors**, but analyzing residuals will essentially lay the wrong answer.

StanKorea

Case study 2 : Marriage age, rate and Divorce rate

Ch. 5
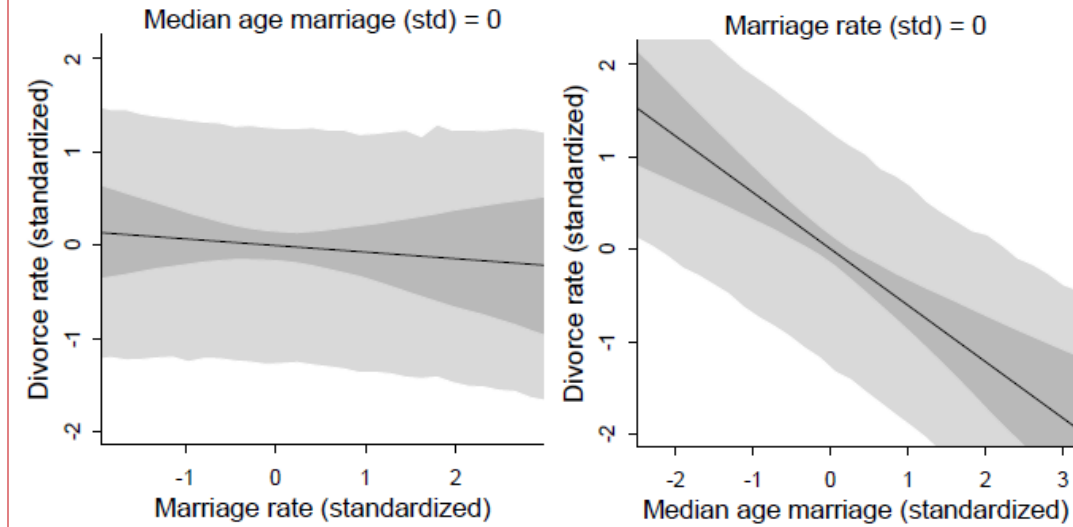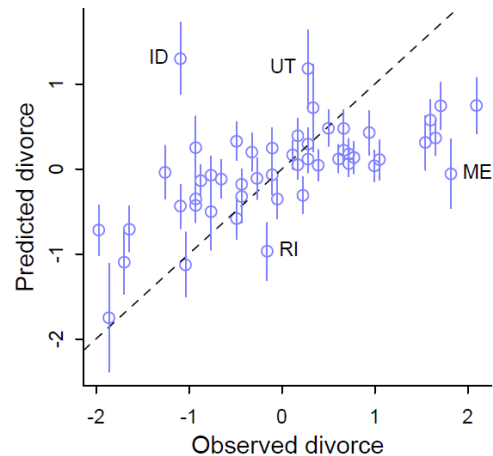
5. Posterior Predictions Plotting

5) The **counterfactual plots** hold other predictors constant and regress one predictor to the outcome. It is very **useful for understanding the models**, but **does not give us the causal inference**. It is plotted on the top.

6) The **posterior prediction plots** are somewhat different with the models before. It compares the posterior distribution of outcome from predictors with raw data, checking the model's accuracy. It is useful to **find model failures**. As the slope of the graph downside goes to unity, the model fits well.

StanKorea
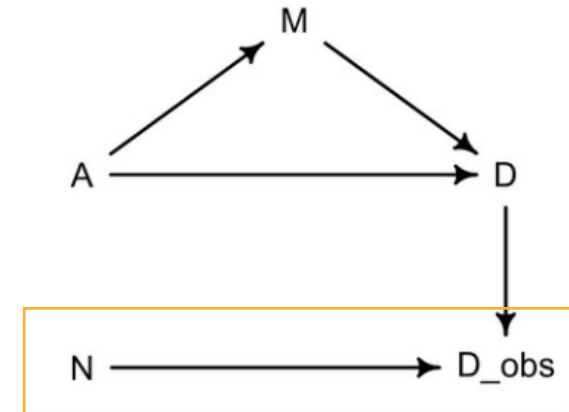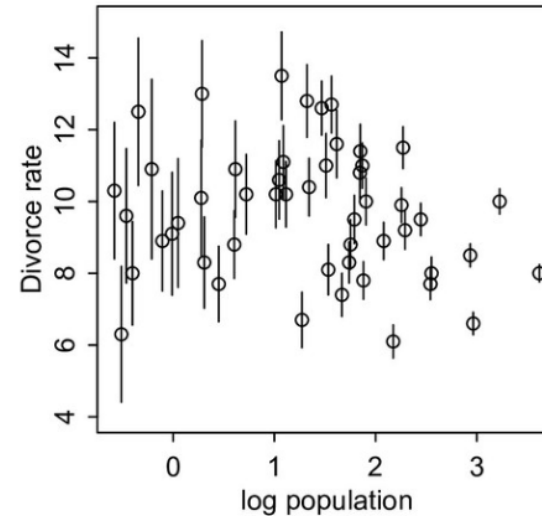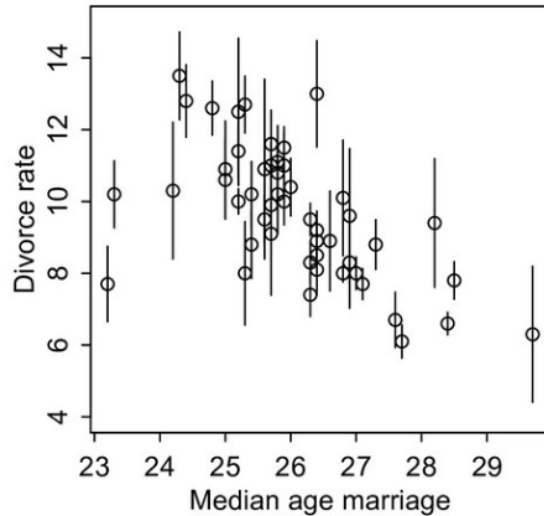
- Relationship between divorce rate, marriage rate and median marriage age for each states

- As for the small state, the standard deviation(line segment) increases due to lack of evidence

- Considering the measurement error, we have to add the D_obs which states the observed divorce rate, connected with the population size of state(N) and the unknown true divorce rate D in the DAG.
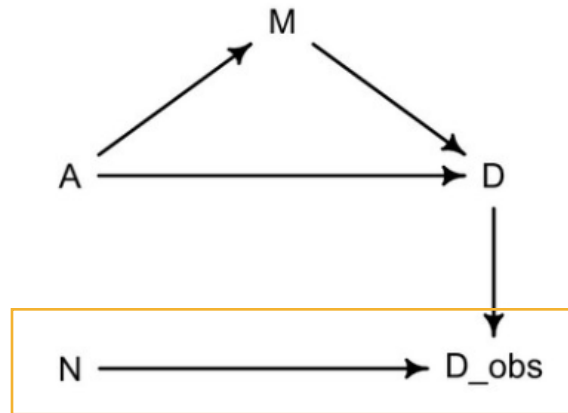
Case study 2 : Marriage age, rate and Divorce rate

Ch. 15

$$D_{\mathrm{OBS},i} \sim \mathrm{Normal}(D_{\mathrm{TRUE},i}, D_{\mathrm{SE},i})$$
$$D_{\mathrm{TRUE},i} \sim \mathrm{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_A A_i + \beta_M M_i$$
$$\alpha \sim \mathrm{Normal}(0, 0.2)$$
$$\beta_A \sim \mathrm{Normal}(0, 0.5)$$
$$\beta_M \sim \mathrm{Normal}(0, 0.5)$$
$$\sigma \sim \mathrm{Exponential}(1)$$

m15.1

- Keep the $D_{\mathrm{TRUE},i}$ for unknown parameter, and assuming that the observed divorce rate ($D_{\mathrm{OBS},i}$) follows the normal distribution having D_{TRUE} as mean value.

- D_{SE} is the standard error of observation, inversely proportional to the size of population.

- A : median marriage age
  M : Marriage rate

Case study 2 : Marriage age, rate and Divorce rate
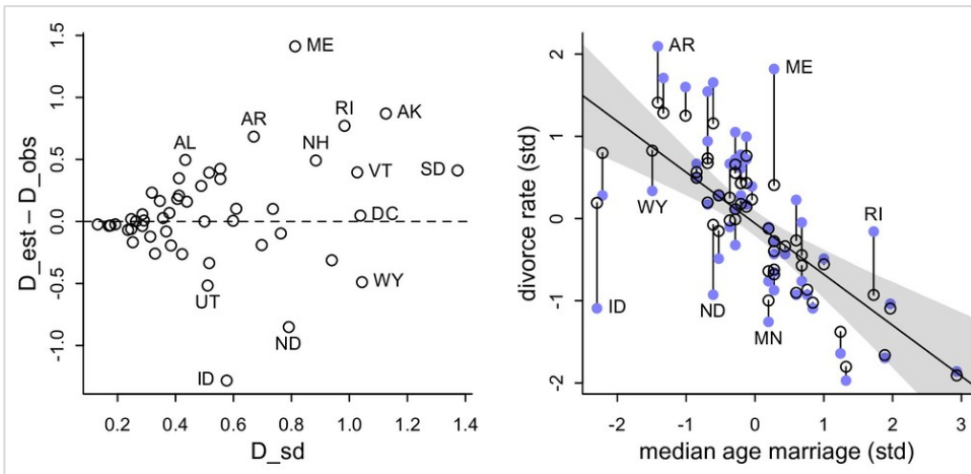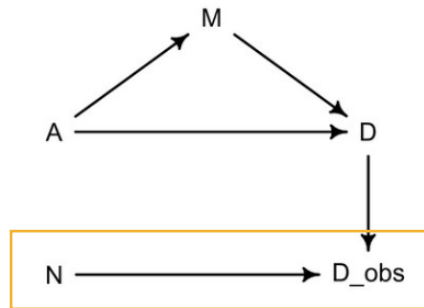
- As in the result plot, the shrinkage happens to the regression model. (blue circle : observed data; open circle : posterior of true data)

- The range of shrinkage has the tendency that **as the standard deviation of the divorce rate(or the median age) increases, the shrinkage becomes more extreme**.

- Such tendency comes from **the population of state**: with lots of evidence, like in California, the observed rate and estimated rate has small difference while the small states like *ID* or *ND* in the figure shrinkages far more because of lack of evidence.

StanKorea

Case study 2 : Marriage age, rate and Divorce rate

Ch. 15
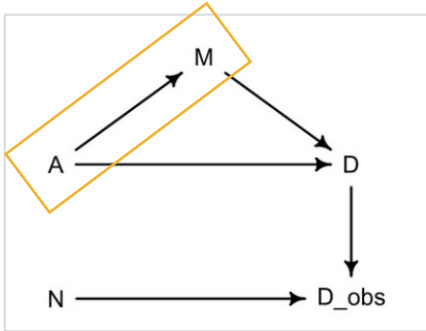
$$
\begin{aligned}
D_{\text{OBS},i} &\sim \text{Normal}(D_{\text{TRUE},i}, D_{\text{SE},i}) \\
D_{\text{TRUE},i} &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_A A_i + \beta_M M_{\text{TRUE},i} \\
M_{\text{OBS},i} &\sim \text{Normal}(M_{\text{TRUE},i}, M_{\text{SE},i}) \\
M_{\text{TRUE},i} &\sim \text{Normal}(0, 1) \\
\alpha &\sim \text{Normal}(0, 0.2) \\
\beta_A &\sim \text{Normal}(0, 0.5) \\
\beta_M &\sim \text{Normal}(0, 0.5) \\
\sigma &\sim \text{Exponential}(1)
\end{aligned}
$$

m15.2

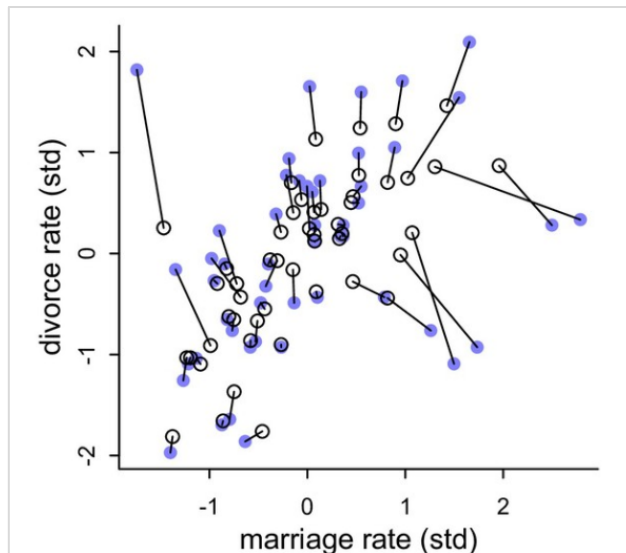- With just the same approach, the error on predictor can also be expressed in Stan code. Let's consider the error in marriage rate.

- As seen in the figure in bottom, the shrinkage also happens but in both horizontal and vertical direction. The point is that, although the fact that the low standard deviation makes lower shrinkage is still valid, **the shrinkage rate of divorce rate and marriage rate is different**.

- Divorce rate shrinks more, because the marriage rate is not strongly associated with the outcome(divorce rate). Although it is the predictor and connected with the outcome in DAG, the low shrinkage means **the marriage rate does not give much information through regression**.

- Also the model is not perfect because in this case the interaction and association between age and marriage rate is neglected.

StanKorea

Case study 3 : Oceanic Tool Complexity
Ch. 11

culture	population	contact	total_tools	mean_TU
Malekula	1100	low	13	3.2
Tikopia	1500	low	22	4.7
Santa Cruz	3600	low	24	4.0
Yap	4791	high	43	5.0
Lau Fiji	7400	high	33	5.0
Trobriand	8000	high	19	4.0
Chuuk	9200	high	40	3.8
Manus	13000	low	28	6.6
Tonga	17500	high	55	5.4
Hawaii	275000	low	71	6.6

1.   Experiment

As seen in the chart, the population, contact to other islands and continents, the number of tools are searched.

2.   Purpose

1) We want to see the **relationship between the complexity of toolkit and the magnitude of population**(log of population).

2) We want to see whether **the contact with other islands moderates impact**. If the island has high contact rate, it can possess lots of tools despite of small population size.

StanKorea
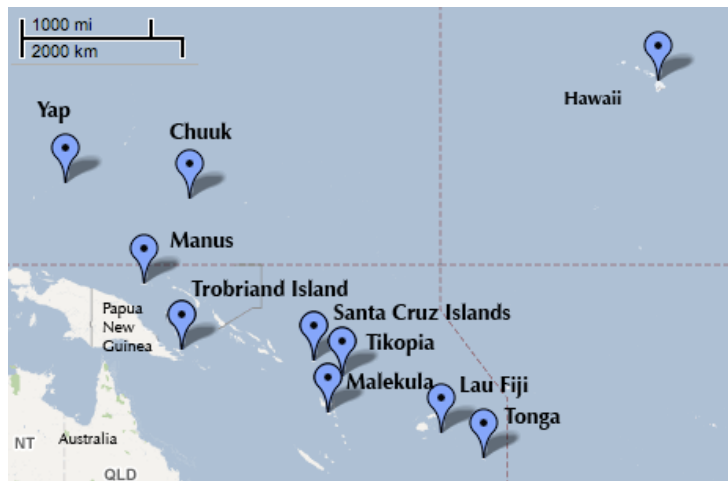
Case study 3 : Oceanic Tool Complexity

Ch. 11

$$T_i \sim \text{Poisson}(\lambda_i)$$

$$\log \lambda_i = \alpha_{\text{CID}[i]} + \beta_{\text{CID}[i]} \log P_i$$

$$\alpha_j \sim \text{Normal}(3, 0.5)$$

$$\beta_j \sim \text{Normal}(0, 0.2)$$

m11.10

```
dat <- list(
    T = d$total_tools ,
    P = d$P ,
    cid = d$contact_id )

# intercept only
m11.9 <- ulam(
    alist(
        T ~ dpois( lambda ),
        log(lambda) <- a,
        a ~ dnorm(3,0.5)
    ), data=dat , chains=4 , log_lik=TRUE )

# interaction model
m11.10 <- ulam(
    alist(
        T ~ dpois( lambda ),
        log(lambda) <- a[cid] + b[cid]*P,
        a[cid] ~ dnorm( 3 , 0.5 ),
        b[cid] ~ dnorm( 0 , 0.2 )
    ), data=dat , chains=4 , log_lik=TRUE )

compare( m11.9 , m11.10 , func=LOO )
```

```
         LOO pLOO dLOO weight    SE   dSE
m11.10  85.5  7.1  0.0      1 13.22    NA
m11.9  141.1  8.0 55.5      0 33.33 32.78
```

3. Method : Poisson GLM

1) **The tool complexity is modeled with Poisson GLM**, one of the counting distribution models, which is very handy since the rate for each trial is very low although the maximum count value is unknown.

2) T_i, λ_i, P_i each means the total number of tools, expected tool numbers and population size for case i. α_{CID} and β_{CID} each indicates the intercept and slope of the model with the CID meaning contact ID. Remember that our final purpose is **to check the relationship between the number of tools and population magnitude.**

3) The code below indicates that the model with both intercept and slope shows lower pLOO, indicating less overfitting. We will use m11.10 model from now on.

StanKorea
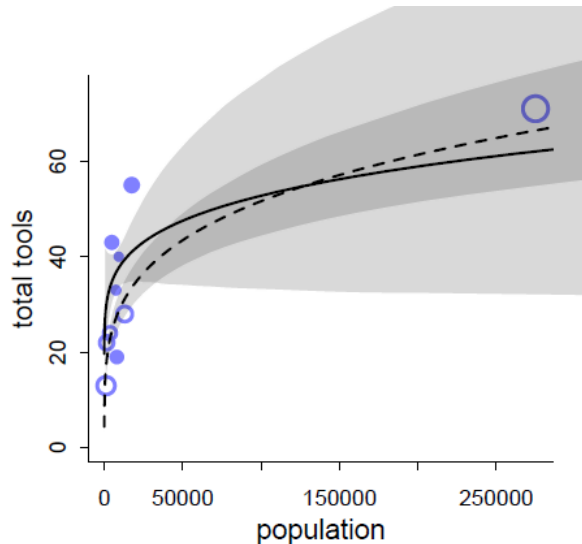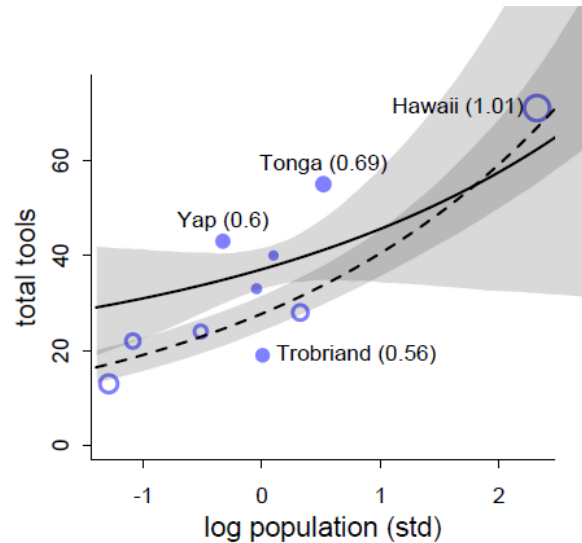
4. Result

1) The chart upper left shows the relationship between log population and total tools, which we want at first. The bottom one represents it with the real population number.

2) The solid line and filled circle show the high-contact islands while the dashed ones show low-contact islands.

3) There are **two main flaws** in the graph based on statistical model:
   - **The lines do not pass the origin**. Zero population should indicate zero tools.
   - As higher the population, **low-contact islands have more tools than high-contact islands.** In common sense, there should be explicit difference.

   Such flaws come from the outfitting case of Hawaii: it has too many population but geographically has low contact.

4) The listed errors show the **limitation of statistical model,** emphasizing the necessity of introducing **scientific model instead.**
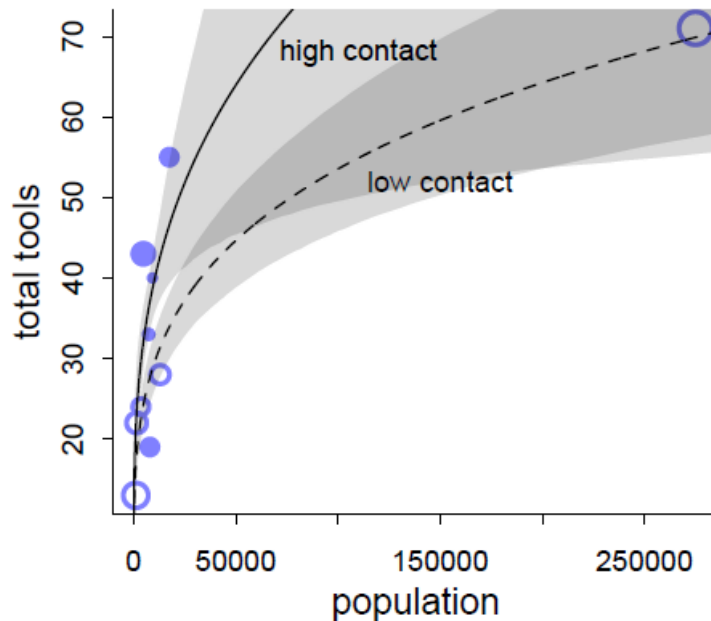
StanKorea

$$\Delta T = \alpha P^{\beta} - \gamma T$$

$$T_i \sim \text{Poisson}(\lambda_i)$$

$$\lambda_i = \alpha P_i^{\beta} / \gamma$$

4. Result - Improved

1) The introduced scientific model is about the rate of number of tools change with time. α and β each shows the innovation rate and diminishing return. As the population increase, the ΔT should increase. But economically as the population reaches to certain level, the **saturation effect occur and each additional person contributes less to the tool increase.** That's the diminishing return effect. γ indicates the loss rate, showing that as the tool number increases, the loss rate also increases.

2) At the steady state, the λ_i is determined as left, which did not include any ad hoc link functions. With the new scientific model, the bottom left lines are drawn. Since the values are scientifically derived, **the two flaws in the statistical models are clearly solved.** It passes origin and shows clear difference.

3) As the result, although it contains some flaws too, **scientific model can be more useful than statistical model in some case** where some outliners exist in the data.

StanKorea

Case study 3 : Oceanic Tool Complexity

Ch.14

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\lambda_i = \alpha P_i^{\beta} / \gamma$$

$\longrightarrow$

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\lambda_i = \exp(k_{\text{SOCIETY}[i]}) \alpha P_i^{\beta} / \gamma$$

expected tools

Varying factor

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\lambda_i = \exp(k_{\text{SOCIETY}[i]}) \alpha P_i^{\beta} / \gamma$$
$$\mathbf{k} \sim \text{MVNormal}\big((0, \ldots, 0), \mathbf{K}\big)$$
$$\mathbf{K}_{ij} = \eta^2 \exp(-\rho^2 D_{ij}^2) + \delta_{ij}(0.01)$$
$$\alpha \sim \text{Exponential}(1)$$
$$\beta \sim \text{Exponential}(1)$$
$$\eta^2 \sim \text{Exponential}(2)$$
$$\rho^2 \sim \text{Exponential}(0.5)$$

m14.7

5. New method

1) We added *varying intercepts* term to our model.

2) Varying intercepts are estimated in light of **geographic distance**. Not distinct category membership like in typical varying intercepts but as a form of correlation matrix.
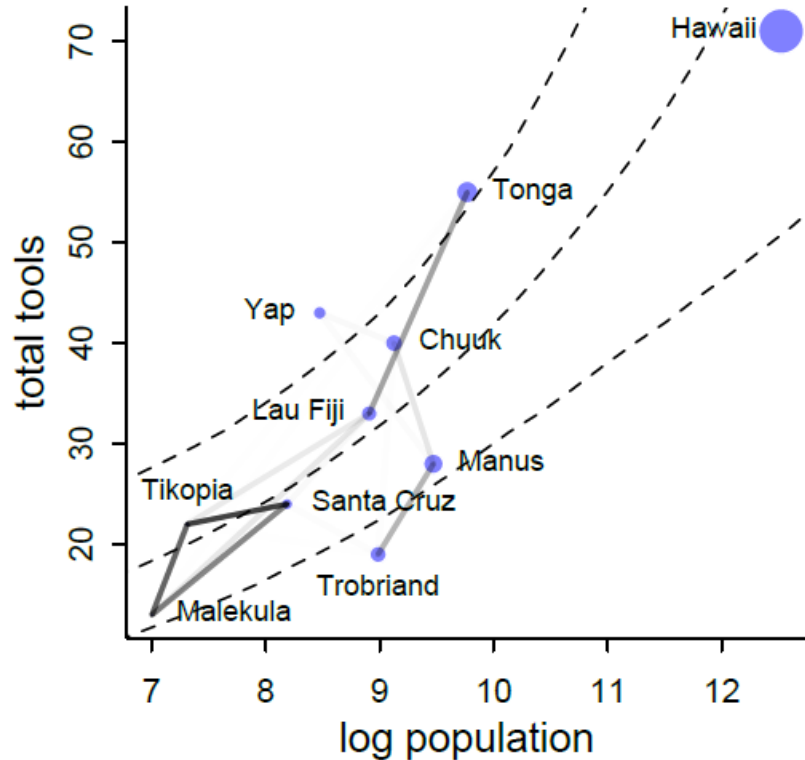
6. New result

1) Graph on the left shows the results. The size of each dots indicate population of the island. Darker lines indicate stronger correlations, with pure white being zero correlation and pure black 100% correlation.

2) Looking at the correlations among Malekula, Tikopia, and Santa Cruz describe the fact that they are below the expected number of tools for their populations. All three societies lying below the expectation and being so close, is **consistent with spatial covariance**.

3) Tonga has more tools than expected for its population, and its proximity to Fiji counteracts some of the tug Fiji's smaller neighbors—Malekula, Tikopia, and Santa Cruz—exert on it. So the model seems to think Fiji would have fewer tools, if it weren't for Tonga.

4) Of course the correlations that this model describes by geographic distance may be the of other, unmeasured commonalities between geographically close societies.

StanKorea

Appendix : Model Codes

Chimpanzee models

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \beta_{\text{TREATMENT}[i]}$$
$$\alpha_j \sim \text{to be determined}$$
$$\beta_k \sim \text{to be determined}$$

m11.4

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \beta_{\text{SIDE}[i]} + \beta_{\text{COND}[i]}$$
$$\alpha_j \sim \text{to be determined}$$
$$\beta_k \sim \text{to be determined}$$

m11.5

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \alpha_{\text{ACTOR}[i]} + \gamma_{\text{BLOCK}[i]} + \beta_{\text{TREATMENT}[i]}$$
$$\beta_j \sim \text{Normal}(0, 0.5) \quad , \text{for } j = 1..4$$
$$\alpha_j \sim \text{Normal}(\bar{\alpha}, \sigma_\alpha) \quad , \text{for } j = 1..7$$
$$\gamma_j \sim \text{Normal}(0, \sigma_\gamma) \quad , \text{for } j = 1..6$$
$$\bar{\alpha} \sim \text{Normal}(0, 1.5)$$
$$\sigma_\alpha \sim \text{Exponential}(1)$$
$$\sigma_\gamma \sim \text{Exponential}(1) \qquad \text{varying intercepts on block}$$

m13.6

$$L_i \sim \text{Binomial}(1, p_i)$$
$$\text{logit}(p_i) = \gamma_{\text{TID}[i]} + \alpha_{\text{ACTOR}[i],\text{TID}[i]} + \beta_{\text{BLOCK}[i],\text{TID}[i]}$$

Mean treatment effects *Each actor in each treatment* *Each block in each treatment*

m14.2

StanKorea

Appendix : Model Codes

Divorce rate models

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_A A_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m5.1

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_M M_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m5.2

$$D_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_M M_i + \beta_A A_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m5.3

$$D_{\text{OBS},i} \sim \text{Normal}(D_{\text{TRUE},i}, D_{\text{SE},i})$$
$$D_{\text{TRUE},i} \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_A A_i + \beta_M M_i$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m15.1

$$D_{\text{OBS},i} \sim \text{Normal}(D_{\text{TRUE},i}, D_{\text{SE},i})$$
$$D_{\text{TRUE},i} \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_A A_i + \beta_M M_{\text{TRUE},i}$$
$$M_{\text{OBS},i} \sim \text{Normal}(M_{\text{TRUE},i}, M_{\text{SE},i})$$
$$M_{\text{TRUE},i} \sim \text{Normal}(0, 1)$$
$$\alpha \sim \text{Normal}(0, 0.2)$$
$$\beta_A \sim \text{Normal}(0, 0.5)$$
$$\beta_M \sim \text{Normal}(0, 0.5)$$
$$\sigma \sim \text{Exponential}(1)$$

m15.2

StanKorea

Appendix : Model Codes

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\log \lambda_i = \alpha_{\text{CID}[i]} + \beta_{\text{CID}[i]} \log P_i$$
$$\alpha_j \sim \text{Normal}\,(3, 0.5)$$
$$\beta_j \sim \text{Normal}\,(0, 0.2)$$

m11.10

$$\Delta T = \alpha P^\beta - \gamma T$$

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\lambda_i = \alpha P_i^\beta / \gamma$$

Scientific Model

$$T_i \sim \text{Poisson}(\lambda_i)$$
$$\lambda_i = \exp(k_{\text{SOCIETY}[i]})\alpha P_i^\beta / \gamma$$
$$\mathbf{k} \sim \text{MVNormal}\big((0, \ldots, 0), \mathbf{K}\big)$$
$$\mathbf{K}_{ij} = \eta^2 \exp(-\rho^2 D_{ij}^2) + \delta_{ij}(0.01)$$
$$\alpha \sim \text{Exponential}(1)$$
$$\beta \sim \text{Exponential}(1)$$
$$\eta^2 \sim \text{Exponential}(2)$$
$$\rho^2 \sim \text{Exponential}(0.5)$$

m14.7

StanKorea

Thank You