Homework 6

Yuki Joyama

2024-03-24

1

$Y_{ij} = \mu + b_i + e_{ij}$ $\quad i = 1, \cdots, m$, $j = 1 \cdots n$, $b_i \sim N(0, \sigma_b^2)$, $e_{ij} \sim N(0, \sigma_e^2)$

b_i and e_{ij} are statistically independent for each i and j —①

e_{ij} and e_{ik} are statistically independent for any two values $j, k = 1, \cdots n$, $j \neq k$ —②

? variance of Y_{ij}, covariance and correlation between any values Y_{ij} and Y_{ik} $(j \neq k)$

$\text{var}(Y_{ij}) = \text{var}(\mu + b_i + e_{ij}) = \sigma_b^2 + \sigma_e^2$ (Given①)

$\text{Cov}(Y_{ij}, Y_{ik}) = \text{Cov}(\mu + b_i + e_{ij}, \mu + b_i + e_{ik})$

$= \text{Cov}(\mu, \mu) + \text{Cov}(\mu, b_i) + \text{cov}(\mu, e_{ik}) + \text{Cov}(b_i, \mu) + \text{Cov}(b_i, b_i)$

$+ \text{Cov}(b_i, e_{ik}) + \text{Cov}(e_{ij}, \mu) + \text{Cov}(e_{ij}, b_i) + \text{Cov}(e_{ij}, e_{ik})$

$= \text{Var}(b_i) = \sigma_b^2$ (Given① and②)

$\text{Corr}(Y_{ij}, Y_{ik}) = \frac{\text{Cov}(Y_{ij}, Y_{ik})}{\sqrt{\text{Var}(Y_{ij})}\sqrt{\text{Var}(Y_{ik})}} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}$

This is compound symmetric structure.

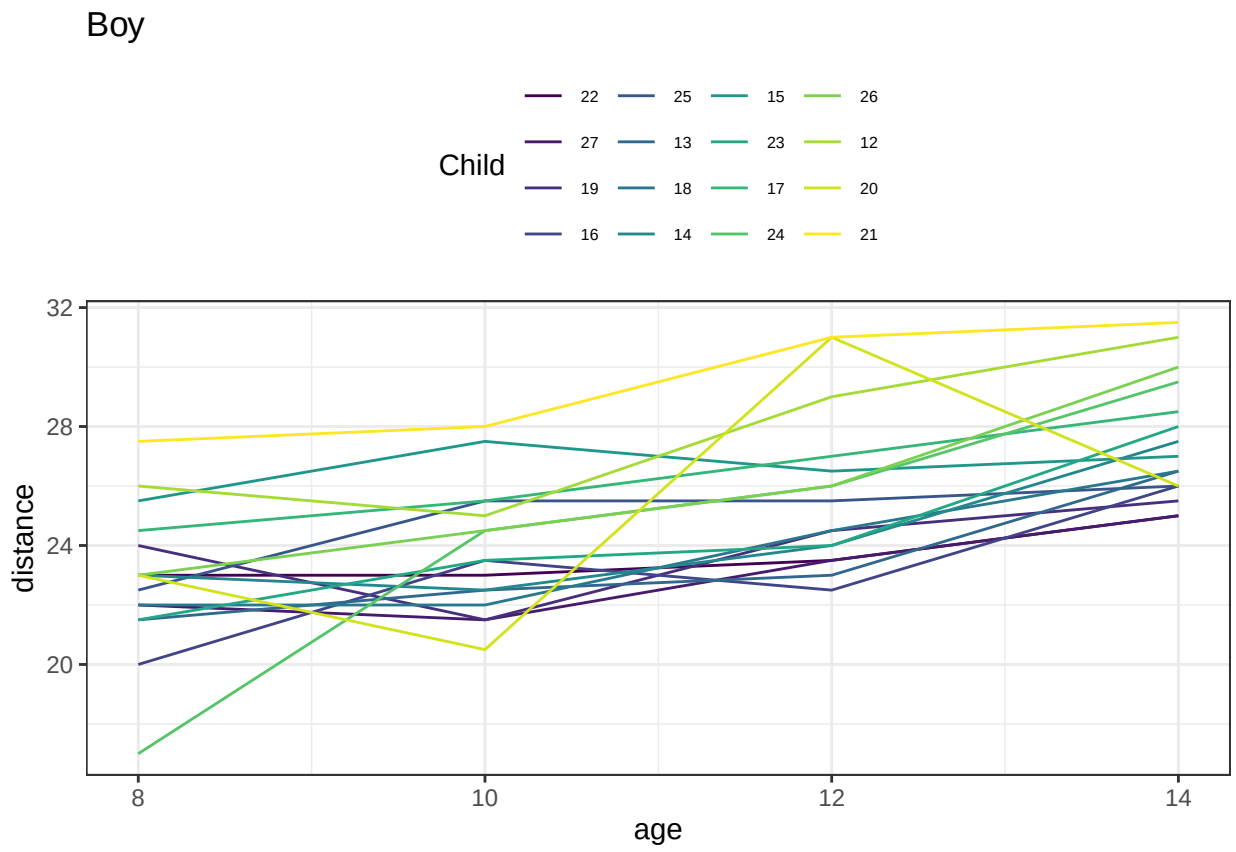
2

```
# data prep
df = read.table("/Users/yukijoyama/Library/CloudStorage/GoogleDrive-jikeyu1995@gmail.com/My Drive/versi
  mutate(
    Gender = as.factor(
      case_when(
        Gender == 0 ~ "girl",
        Gender == 1 ~ "boy"
      )
    )
  )
```

a

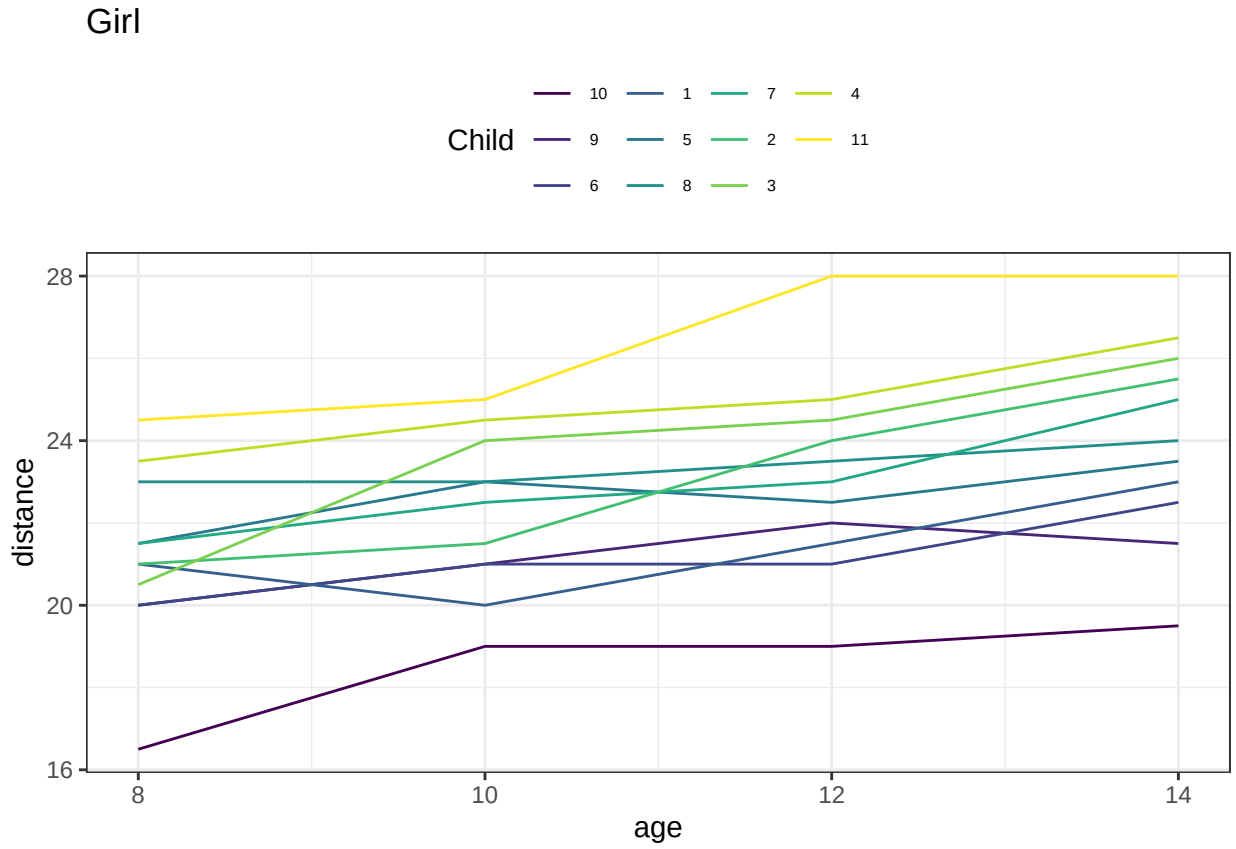
```
# grouped data
df_new = groupedData (Distance ~ Age | Child, data = as.data.frame(df))

# create a spaghetti plot
# boy
df_new |>
  filter(Gender == "boy") |>
  ggplot(aes(x = Age, y = Distance, group = Child, color = Child)) +
  geom_line() + # spaghetti plot
  theme(legend.text = element_text(size = 6)) + # changed legend text size
  labs(
    title = "Boy",
    x = "age",
    y = "distance"
  ) +
  viridis::scale_color_viridis(
    discrete = TRUE
  )
```




```
# girl
df_new |>
  filter(Gender == "girl") |>
  ggplot(aes(x = Age, y = Distance, group = Child, color = Child)) +
```

```
geom_line() + # spaghetti plot
theme(legend.text = element_text(size = 6)) + # changed legend text size
labs(
  title = "Girl",
  x = "age",
  y = "distance"
) +
viridis::scale_color_viridis(
  discrete = TRUE
)
```



Distance tends to increase with age, with boys having relatively higher distance values than girls.

b

The marginal form is

$$E(Y_{ij}) = \beta_0 + \beta_1 age_{ij}$$

c

Compound symmetry covariance

```
# summary(gls(Distance ~ Gender + Age, data = df, correlation = corCompSymm(form = ~ 1 | Gender), metho
```

Exponential covariance

Autoregressive covariance