HW6

Siyan Chen

4/1/2019

PROBLEM1

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

$cov(Y_{ij}, Y_{ik}) = E[(Y_{ij} - \mu_{ij})(Y_{ik} - \mu_{ik})] = E(b_i + e_{ij})(b_i + e_{ik}) = E({b_i}^2 + e_{ij} * e_{ik} + b_i * e_{ik} + b_i * e_{ij})$
$= E({b_i}^2) + E(e_{ij} * e_{ik}) + E(b_i * e_{ik}) + E(b_i * e_{ij}) = E(b_i^2) = (Eb_i)^2 + var(b_i) = \sigma_b^2$

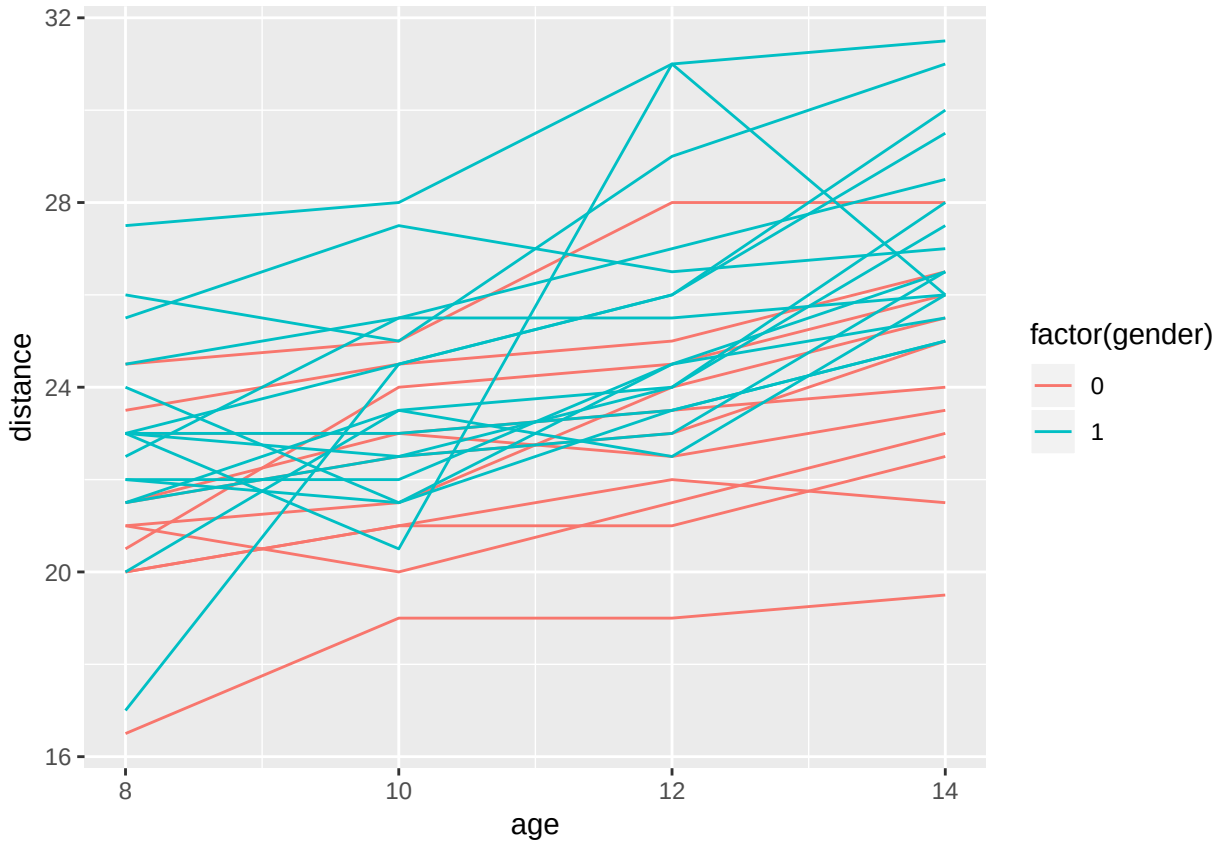
$corr(Y_{ij}, Y_{ik}) = \frac{cov(Y_{ij}, Y_{ik})}{\sqrt{var(Y_{ij}) * var(Y_{ik})}}$

$= {\sigma_b}^2/({\sigma_b}^2 + {\sigma_e}^2)$

PROBLEM 2

2.1

```
ggplot(dental_df, aes(x = age, y = distance, group = child)) + geom_line(aes(color = factor(gender)))
```



2.2

$E(Y_{ij}) = \beta_0 + \beta_1 * age_{ij}$
$var(Y_{ij}) = var(a_i) + var(b_k) + var(e_{ij}) = {\sigma_a}^2 + {\sigma_b}^2 + {\sigma_e}^2$

2.3

```
### compound symmetry
comsym = gls(distance ~ age + gender,dental_df, correlation=corCompSymm(form = ~ 1|child),weights=varId
summary(comsym)
```

```
## Generalized least squares fit by REML
##   Model: distance ~ age + gender
##   Data: dental_df
##        AIC      BIC    logLik
##   452.1147 473.3464 -218.0574
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | child
##  Parameter estimate(s):
##       Rho
## 0.6167736
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | age
##  Parameter estimates:
##         8        10        12        14
## 1.0000000 0.8745886 1.0420364 0.9649285
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept) 15.395761 0.8814993 17.465425  0.0000
## age          0.663901 0.0617772 10.746694  0.0000
## gender       2.151163 0.7450111  2.887424  0.0047
##
##  Correlation:
##        (Intr) age
## age    -0.759
## gender -0.501  0.000
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.47252111 -0.57661430  0.02505237  0.58013176  2.22216086
##
## Residual standard error: 2.369293
## Degrees of freedom: 108 total; 105 residual
```

```
corMatrix(comsym$modelStruct$corStruct)[[1]]
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.6167736 0.6167736 0.6167736
## [2,] 0.6167736 1.0000000 0.6167736 0.6167736
## [3,] 0.6167736 0.6167736 1.0000000 0.6167736
## [4,] 0.6167736 0.6167736 0.6167736 1.0000000
```

```r
### exponential
exp = gls(distance ~ age + gender,dental_df, correlation=corExp(form = ~ 1|child), method="REML")
summary(exp)
```

```
## Generalized least squares fit by REML
##   Model: distance ~ age + gender
##   Data: dental_df
##        AIC      BIC    logLik
##   455.4483 468.7181 -222.7241
##
## Correlation Structure: Exponential spatial correlation
##  Formula: ~1 | child
##  Parameter estimate(s):
##    range
## 2.133938
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept) 15.459995 1.1309319 13.670138   0e+00
## age          0.652960 0.0906420  7.203723   0e+00
## gender       2.418714 0.6933441  3.488476   7e-04
##
##  Correlation:
##        (Intr) age
## age    -0.882
## gender -0.363  0.000
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.65148775 -0.69592567 -0.06214639  0.48659340  2.29666951
##
## Residual standard error: 2.301495
## Degrees of freedom: 108 total; 105 residual
```

```r
corMatrix(exp$modelStruct$corStruct)[[1]]
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.6258671 0.3917097 0.2451582
## [2,] 0.6258671 1.0000000 0.6258671 0.3917097
## [3,] 0.3917097 0.6258671 1.0000000 0.6258671
## [4,] 0.2451582 0.3917097 0.6258671 1.0000000
```

```r
### autoregressive
auto1 = gls(distance ~ age + gender,dental_df, correlation=corAR1(form = ~ 1|child), method="REML")
summary(auto1)
```

```
## Generalized least squares fit by REML
##   Model: distance ~ age + gender
##   Data: dental_df
##        AIC      BIC    logLik
##   455.4483 468.7181 -222.7241
##
## Correlation Structure: AR(1)
##  Formula: ~1 | child
##  Parameter estimate(s):
```

```
##       Phi
## 0.6258671
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept) 15.459995 1.1309319 13.670138   0e+00
## age          0.652960 0.0906420  7.203723   0e+00
## gender       2.418714 0.6933441  3.488476   7e-04
##
##  Correlation:
##        (Intr) age
## age    -0.882
## gender -0.363  0.000
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.65148770 -0.69592566 -0.06214639  0.48659339  2.29666947
##
## Residual standard error: 2.301495
## Degrees of freedom: 108 total; 105 residual
```

```
corMatrix(auto1$modelStruct$corStruct)[[1]]
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.6258671 0.3917097 0.2451582
## [2,] 0.6258671 1.0000000 0.6258671 0.3917097
## [3,] 0.3917097 0.6258671 1.0000000 0.6258671
## [4,] 0.2451582 0.3917097 0.6258671 1.0000000
```

The three method produce similar coefficient estimates. For model with compound symmetry covariance, variance is constant and correlation between any two visits is 0.617. Exponential and autogressive covariance have same covariance and coefficient estimae.Exponential covariance is a gneneralization of AR(1) and correlation decreases exponentially as age difference increases. It reduces to AR(1) when all response time are same and the correlation decreases between two ages.