

2132 HW9

Orly Olbum

due April 22, 2021

Problem 1

This problem is an extension of Problem 4 from HW8.

#(a) Use REML + GLS to fit the model you proposed in part (c)(i) of Problem 4. Test the null hypothesis that there is no patient x treatment interaction. Make sure to report your null and alternative models.

From HW8 #4 (c)(i): The null model, with no interaction:

$$y_{ij} = \mu_{..} + \rho_i + \tau_j + \epsilon_{ij}$$

The alternative model, including interaction:

Hours = mean + Individual (random) + Treatment (fixed) + Individual*Treatment + error

$$y_{ij} = \mu_{..} + \rho_i + \tau_j + (\rho\tau)_{ij} + \epsilon_{ij}$$

```
mod1_int = lmer(Hours ~ Treatment + Treatment:Individual + (1|Individual), data = sleep)
mod1_none = lmer(Hours ~ Treatment + (1|Individual), data = sleep)
anova(mod1_int, mod1_none, refit = FALSE)
```

```
## Data: sleep
## Models:
## mod1_none: Hours ~ Treatment + (1 | Individual)
## mod1_int: Hours ~ Treatment + Treatment:Individual + (1 | Individual)
##          npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## mod1_none     6 317.91 328.05 -152.96   305.91
## mod1_int     10 303.41 320.30 -141.71   283.42 22.499   4  0.0001594 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Because ANOVA produces a statistically significant p-value, we reject the null that the interaction term doesn't apply and we conclude that the interaction term is significant.

#(b) Use the model you fitted above to estimate the parameters in the model you proposed part (c)(ii) of Problem 4. Which drug would you recommend an insomniac take if they slept an average of 2 hours a night without treatment? How about one that slept an average of 6 hours a night?

Model from (c)(ii):

Hours|y0 = mean + Individual (random) + Treatment (fixed) Individual*Treatment + error

$$y_{ij}|y0 = y0 + \rho_i + \tau_j + (\rho\tau)_{ij} + \epsilon_{ij}$$

If we use the model from above to fit the model with the mean now as intercept, we might recommend Treatment B for someone who has slept an average of 2 hours and Treatment C for someone who has slept an average of 6 hours a night.

$$2. Y = X\beta + \epsilon \quad X \in \mathbb{R}^{n \times p} \text{ is full rank, } \beta \in \mathbb{R}^p$$

$E(\epsilon) = 0$, $\text{var}(\epsilon) = V(\theta)$ for $\theta \in \mathbb{R}^b$ and $V: \mathbb{R}^b \rightarrow \mathbb{R}^{n \times n}$

Assume $V(\theta)$ is pd and let $A_x \in \mathbb{R}^{n \times (n-p)}$ be any matrix whose columns form a basis for $\text{Ker}(X^T)$. A_x is not unique.

$$\textcircled{a} \text{ Let } Q = I_n - X(X^T X)^{-1} X^T$$

$$\text{Show that } A_x (A_x^T A_x)^{-1} A_x^T = Q \rightarrow H$$

$$A_x (A_x^T A_x)^{-1} A_x^T = I_n - X(X^T X)^{-1} X^T$$

Since $\text{col}(A)$ form a basis for $\text{Ker}(X^T)$,

$$A^T X = 0 \text{ and } A A^T = Q, \text{ and } A^T A = I_n$$

$$\rightarrow A_x (A_x^T A_x)^{-1} A_x^T = I_n - X(X^T X)^{-1} X^T$$

(for simplicity $\rightarrow A_x = A$)

$$A (A^T A)^{-1} A^T = A^T A - X(X^T X)^{-1} X^T$$

$$A^T A (A^T A)^{-1} A^T = A A^T A A^T - A^T X (X^T X)^{-1} X^T A$$

$$I_n I_n^{-1} I_n = A I_n A^T - 0 \quad \left. \right\} I - H = Q,$$

$$I_n = A^T A I_n \quad \left. \right\} A = Q^T A = A$$

$I_n = I_n$, which checks out,

$$\text{so } A_x (A_x^T A_x)^{-1} A_x^T = Q$$

\textcircled{b} Let $\tilde{Y} = QY \in \mathbb{R}^n$ be residuals after regressing out X from Y . Find expressions for $E(\tilde{Y})$ and $\text{var}(\tilde{Y})$, show that \tilde{Y} is degenerate dist.

$$\tilde{Y} = QY = QX\beta + QE$$

$$\text{Since } Q = A_x (A_x^T A_x)^{-1} A_x^T, \quad QX = 0 \\ \rightarrow = QE$$

$$E(\tilde{Y}) = E(Q\epsilon) = 0_n$$

$$\text{var}(\tilde{Y}) = \text{var}(QE) = QV(\theta)Q^T \quad (Q^T = Q)$$

\tilde{Y} is degenerate:

Because \tilde{Y} is in the image of

the kernel of X^T (Q is orth. matrix),

then \tilde{Y} has degenerate distribution

(2, continued)

② Define

$$\tilde{L}(\theta) = -\frac{1}{2} \log [\det^+ \{ QV(\theta)Q \}] - \frac{1}{2} \tilde{Y}^T \{ QV(\theta)Q \}^{-1} \tilde{Y}$$

B^+ → Moore-Penrose pseudo inverse of $B \in \mathbb{R}^{n \times n}$ and

$\det^+(B)$ is pseudo determinant \rightarrow calling $Ax = A$ for
 $\downarrow = \lim_{\epsilon \rightarrow 0} \frac{\det(B + I_n)}{\epsilon^{\text{rank}(B)}}$ this problem...

$$(i) \text{ Show that } \{ QV(\theta)Q \}^{-1} = A \times \{ A^T V(\theta) A \}^{-1} A^T$$

We know that $X \in \mathbb{R}^{n \times p}$ and $A \in \mathbb{R}^{n \times (n-p)}$ are full rank

$$A^+ = (A^T A)^{-1} A^T \rightarrow H = X X^+$$

$$X^+ = (X^T X)^{-1} X^T \rightarrow Q = I - X X^+ = A A^+ \text{ (yuoma)}$$

We also know that $A^T X = 0$, so $\rightarrow V = V(\theta)$ for

$$A^T (V^{\frac{1}{2}})^{-1} V^{-\frac{1}{2}} X = 0 \rightarrow (V^{\frac{1}{2}} A)^T V^{-\frac{1}{2}} X = 0 \quad \text{simplicity}$$

because V is symmetric and PSD

$$\text{From } I - H = I - X(X^T X)^{-1} X^T = A(A^T A)^{-1} A^T$$

$$\rightarrow I - V^{\frac{1}{2}} X(X^T V^{-1} X)^{-1} X^T V^{-\frac{1}{2}} = V^{\frac{1}{2}} A(A^T A)^{-1} A^T V^{\frac{1}{2}}$$

$$\rightarrow V^{-1} - V^{-1} X(X^T V^{-1} X)^{-1} X^T V^{-1} = A(A^T A)^{-1} A^T \\ A(A^T V A)^{-1} A^T$$

and we have $Q = I - X X^+ = A A^+$, so need to show

$$\text{that } A(A^T V A)^{-1} A^T = Q(QVQ)^+ Q = (QVQ)^+$$

$$A^T Q V Q A / (A^+ (QVQ)^- (A^+)^T) A^T Q V Q A \} (A^T Q V Q A)^-$$

$$= A^T Q V Q (QVQ)^- Q V Q A \} = A^+ (QVQ)^- (A^+)^T$$

$$= A^T Q V Q A$$

$$V^{-1} - V^{-1} X(X^T V^{-1} X)^{-1} X^T V^{-1} = A(A^T V A)^{-1} A^T$$

$$= Q A (A^T V A)^{-1} A^T Q$$

$$= Q A (A^T Q V Q A)^{-1} A^T Q$$

$$= Q A (A^+ (QVQ)^+ (A^+)^T) A^T Q$$

$$= Q (QVQ)^+ Q$$

$$Q (QVQ)^+ Q = A(A^T V A)^{-1} A^T$$

$$\rightarrow Q (QVQ)^+ Q = Q A (A^T V A)^{-1} A^T Q$$

$$\rightarrow (QVQ)^+ = A(A^T V A)^{-1} A^T$$

which is

$$\{ QV(\theta)Q \}^{-1} = A \times \{ A^T V(\theta) A \}^{-1} A^T$$

(2, continued)

- ② (ii) If $B \in \mathbb{R}^n$ is symmetric psd, show that $\det+(B)$ is the product of non-zero eigenvalues of B . Use to show that $\det+\{QV(\theta)Q\} = \det\{A^T A x\}^{-1} \det\{A x^T V(\theta) A x\}$ (again using $V = V(\theta)$ and $A = Ax$)

For $B \in \mathbb{R}^n$, with eigenvalues λ_i , there are $i=n$ non-negative λ_i , and $\text{rank}(B) = r$ which is # of non-zero λ_i (because B is psd)

So, say $r \lambda_i$ are non-zero, and $n-r$ are zero. $B + \epsilon I_n$ also has n eigenvalues with r such $\lambda_i + \epsilon$, and $(n-r)$ are just ϵ .

$$\begin{aligned}\det+(B) &= \lim_{\epsilon \rightarrow 0} \frac{\det(B + \epsilon I_n)}{\epsilon^{\text{rank}(B)}} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\prod_{i=1}^r (\lambda_i + \epsilon)}{\epsilon^{n-r}} \cdot \epsilon^{n-r} \\ &= \lim_{\epsilon \rightarrow 0} \prod_{i=1}^r (\lambda_i + \epsilon) = \prod_{i=1}^r \lambda_i\end{aligned}$$

which gives us the product of non-zero eigenvalues of B .

then, $\det+(QVQ)$ is the product of non-zero eigenvalues of QVQ

$$(QVQ)^+ = A(A^T V A)^{-1} A^T$$

Since $Q = A(A^T A)^{-1} A^T$,

$$\det+\{QVQ\} = \det+\{A(A^T A)^{-1} A^T V A(A^T A)^{-1} A^T\}$$

and from (i),
$$\begin{aligned}&= \det(A^T A)^{-1} \det(A^T V A) \\ &= \det\{A x^T A x\}^{-1} \det\{A x^T V(\theta) A x\}\end{aligned}$$

(2, continued)

② Define REML objective function:

$$l_{REML}(\theta) = -\frac{1}{2} \log [\det \{A_x^T V(\theta) A_x\}] - \frac{1}{2} (A_x^T Y)^T \{A_x^T V(\theta) A_x\}^{-1} (A_x^T Y)$$

Show that $l_{REML}(\theta) = \tilde{l}(\theta) + C$, some constant that does not depend on θ .

use this to show that

$$\hat{\theta}_{REML} = \underset{\theta \in \Theta}{\operatorname{argmax}} \quad l_{REML}(\theta) = \underset{\theta \in \Theta}{\operatorname{argmax}} \quad \tilde{l}(\theta)$$

why does this imply that $\hat{\theta}_{REML}$ is invariant to the choice of A_x ? ($V = V(\theta)$, $A = A_x$)

$$\begin{aligned} \tilde{l}(\theta) &= -\frac{1}{2} \log [\det_+(QVQ)] - \frac{1}{2} \tilde{Y}^T (QVQ)^+ \tilde{Y} \\ l(\theta)_{REML} &= -\frac{1}{2} \log [\det(A^TVA)] - \frac{1}{2} (A^T Y)^T (A^TVA)^{-1} (A^T Y) \\ &= -\frac{1}{2} \log \left[\frac{\det_+(QVQ)}{\det(A^T A)} \right] - \frac{1}{2} \underbrace{Y^T (QVQ) (A^TVA)^{-1} QVQ}_{\tilde{Y}^T (QVQ)^+ \tilde{Y}} \\ &= -\frac{1}{2} \log [\det_+(QVQ)] + \frac{1}{2} \log [\det(A^T A)^{-1}] \\ &\quad - \frac{1}{2} \tilde{Y}^T (QVQ)^+ \tilde{Y} \\ &= \tilde{l}(\theta) + \underbrace{\frac{1}{2} \log [\det(A^T A)^{-1}]}_{\text{does not depend on } \theta} \end{aligned}$$

Since the C term doesn't depend on θ , $\hat{\theta}_{REML}$ will be the same for both $l_{REML}(\theta)$ and $\tilde{l}(\theta)$ because the derivative takes the C term away.

③ Suppose $\theta = \sigma^2 > 0$, $V(\theta) = \sigma^2 I_n$

Show that $\hat{\sigma}_{REML}^2 = \text{MSE}$, $\text{MSE} = (n-p)^{-1} \tilde{Y}^T \tilde{Y}$

we know that

$$\begin{aligned} \hat{\theta} &= \hat{\sigma}^2 = \underset{\theta \in \Theta}{\operatorname{argmax}} \quad l_{REML}(\theta) = \underset{\theta \in \Theta}{\operatorname{argmax}} \quad \tilde{l}(\theta) \\ &= \underset{\theta \in \Theta}{\operatorname{argmax}} \left[-\frac{1}{2} \log [\det_+(\theta \sigma^2 I_n Q)] - \frac{1}{2} \tilde{Y}^T (\theta \sigma^2 I_n Q)^+ \tilde{Y} \right] \\ \frac{\partial \tilde{l}}{\partial \sigma^2} &= -\frac{1}{2} \operatorname{tr} \left[(\theta \sigma^2 I_n Q)^+ \theta I_n Q \right] - \frac{1}{2\sigma^2} \tilde{Y}^T (Q I_n Q)^+ \tilde{Y} \\ &= -\frac{1}{2\sigma^2} (Q I_n Q)^+ (n-p) - \frac{1}{2\sigma^2} \tilde{Y}^T (\theta I_n Q)^+ \tilde{Y} \end{aligned}$$

Set equal $\frac{n-p}{\sigma^2} = \frac{\tilde{Y}^T \tilde{Y}}{\sigma^4} \rightarrow \hat{\sigma}_{REML}^2 = \frac{1}{n-p} \tilde{Y}^T \tilde{Y} = \text{MSE}$

Problem 3

The data *Age1.txt* and *Age3.txt* contain measurements of the concentration of bilirubin, a ubiquitous small molecule metabolite, in the blood plasma of children at age 1 and 3, respectively. The goal is to understand relationship between the concentration of bilirubin and recurrent wheeze, the latter of which is a diagnosis of ≥ 4 wheezing episodes in that year. The complete list of covariates is given below.

- IndividualID: A unique ID given to each child. There should be 33 children whose bilirubin concentration was measured at both age 1 and age 3. You may assume that data collected on different individuals are independent.
- Bilirubin: The log-concentration of bilirubin. Treat this as the dependent variable.
- Wheeze: A factor variable with 3 levels. level 0: no wheezing episodes in that year; level 1: 1-3 wheezing episodes in that year; level 2: ≥ 4 wheezing episodes in that year.
- Diet: A factor variable with 2 levels. level 0: exclusively breastfed for the first six months of life; level 1: not exclusively breastfed for the first six months of life.
- Daycare: A factor variable with 2 levels. level 0: did not attend daycare in the first year of life; level 1: attended daycare in the first year of life.
- Sex: A factor variable with 2 levels. level 0: male; level 1: female.

#(a) By performing two separate linear regressions (one at age 1 and one at age 3) and treating all of the above-mentioned covariates in your model as additive fixed effects, estimate the expected difference in the log-concentration of bilirubin between recurrent wheezers (those with ≥ 4 wheezing episodes in that year) and healthy controls (those who did not wheeze in that year) at ages 1 and 3. Report 95% confidence interval for both expected differences.

```

fit1 = lm(Bilirubin ~ as.factor(Wheeze) + as.factor(Diet) + as.factor(Daycare) + as.factor(Sex), data = age1)

mu1 = mean(fit1$fitted.values[as.numeric(age1$Wheeze) == 0])
mu2 = mean(fit1$fitted.values[as.numeric(age1$Wheeze) == 2])
t = qtukey(p = 0.95, nmeans = 55, df = 11)/sqrt(2)
n1 = sum(age1$Wheeze == 0)
n2 = sum(age1$Wheeze == 2)
sig = sqrt(get_mse(fit1)*((1/n1) + (1/n2)))
lo = (mu1 - mu2) - t*sig
hi = (mu1 - mu2) + t*sig
ci = c(lo, hi)
ci

## [1] -1.073784  1.745235

fit2 = lm(Bilirubin ~ as.factor(Wheeze) + as.factor(Diet) + as.factor(Daycare) + as.factor(Sex), data = age3)

mu1.2 = mean(fit2$fitted.values[as.numeric(age3$Wheeze) == 0])
mu2.2 = mean(fit2$fitted.values[as.numeric(age3$Wheeze) == 2])
t.2 = qtukey(p = 0.95, nmeans = 109, df = 18)/sqrt(2)
n1.2 = sum(age3$Wheeze == 0)
n2.2 = sum(age3$Wheeze == 2)
sig.2 = sqrt(get_mse(fit2)*((1/n1.2) + (1/n2.2)))

```

```

lo.2 = (mu1.2 - mu2.2) - t.2*sig.2
hi.2 = (mu1.2 - mu2.2) + t.2*sig.2
ci.2 = c(lo.2, hi.2)
ci.2

```

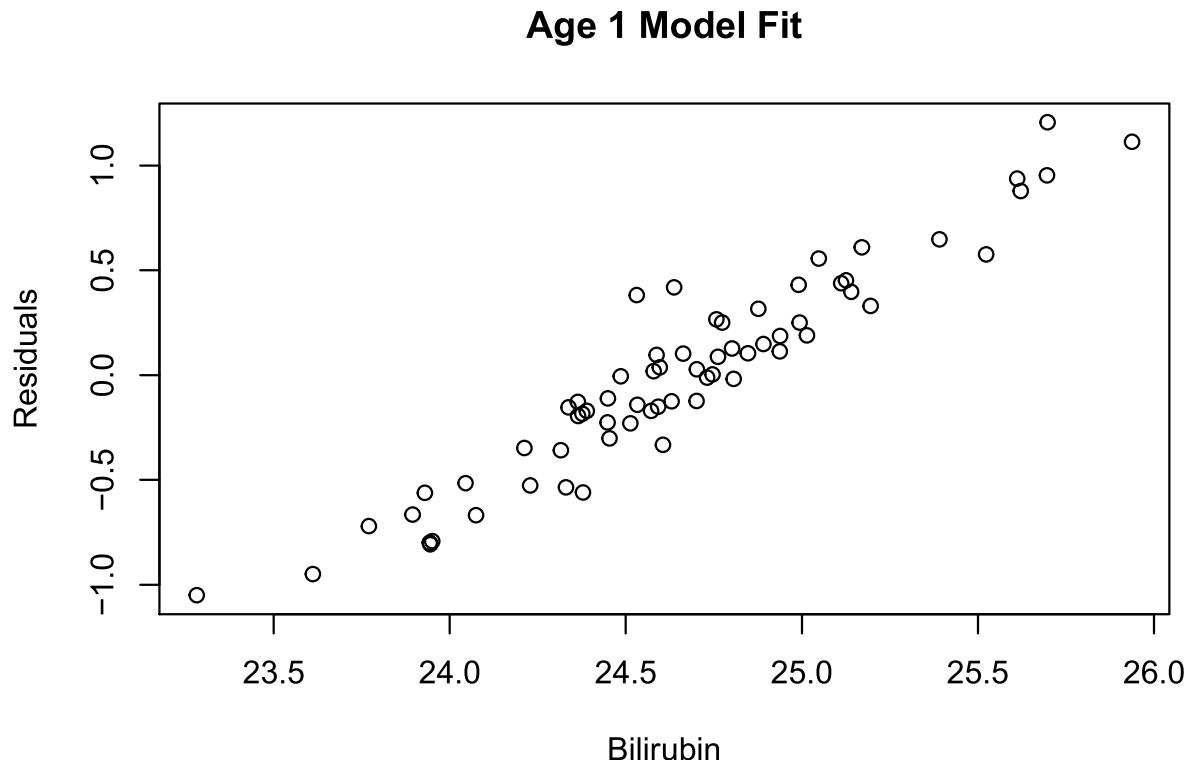
```
## [1] -0.7380337 0.9275604
```

The average difference for age1 is 0.336 and the 95% CI is (-1.074, 1.746); the average difference for age3 is 0.095 and the 95% CI for age3 is (-0.741, 0.931).

#(b) *The inference you performed in part (a) presumably relied on approximating test statistics with a normal or t-distribution. Do you trust this approximation? Give an argument as to why you do or do not. Include plots if necessary.*

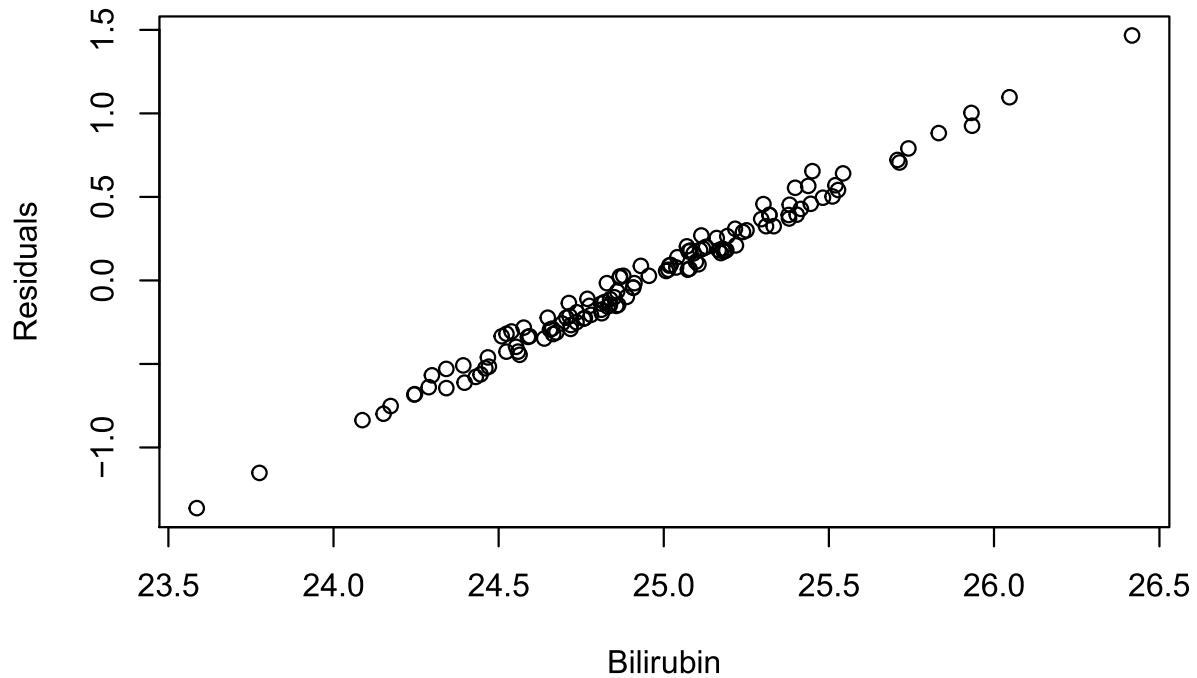
To test this theory we can plot the residuals of the models against the Bilirubin variable and see if there is a straight line.

```
plot(age1$Bilirubin, fit1$residuals, xlab = "Bilirubin", ylab = "Residuals", main = "Age 1 Model Fit")
```



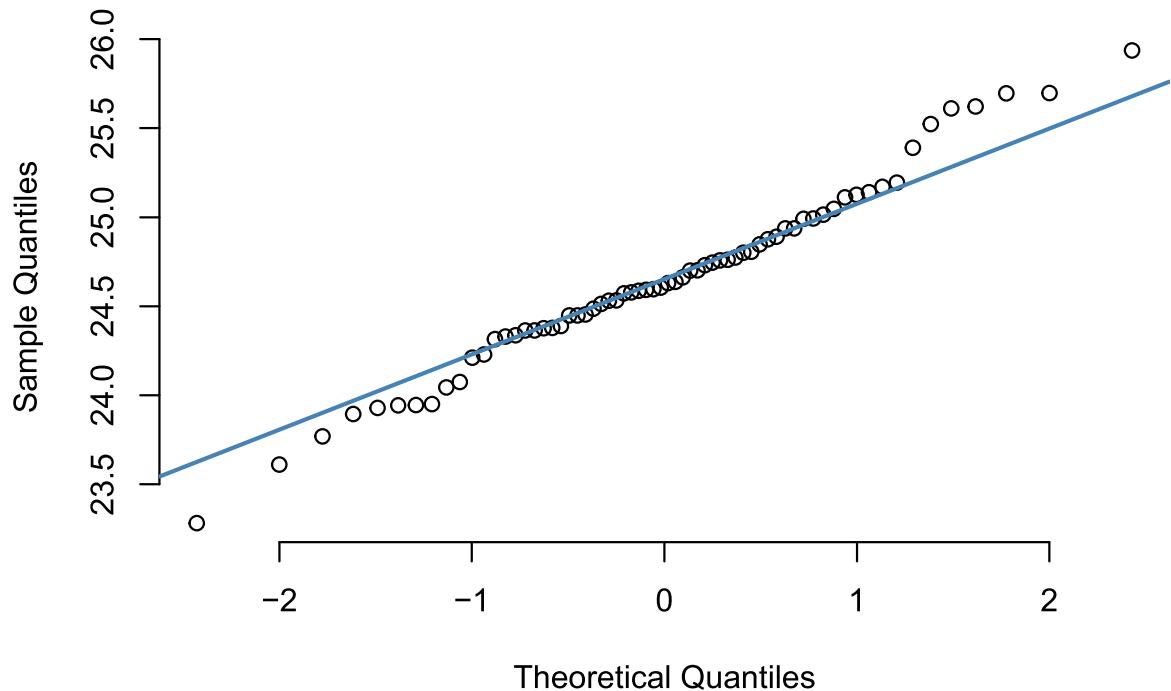
```
plot(age3$Bilirubin, fit2$residuals, xlab = "Bilirubin", ylab = "Residuals", main = "Age 3 Model Fit")
```

Age 3 Model Fit

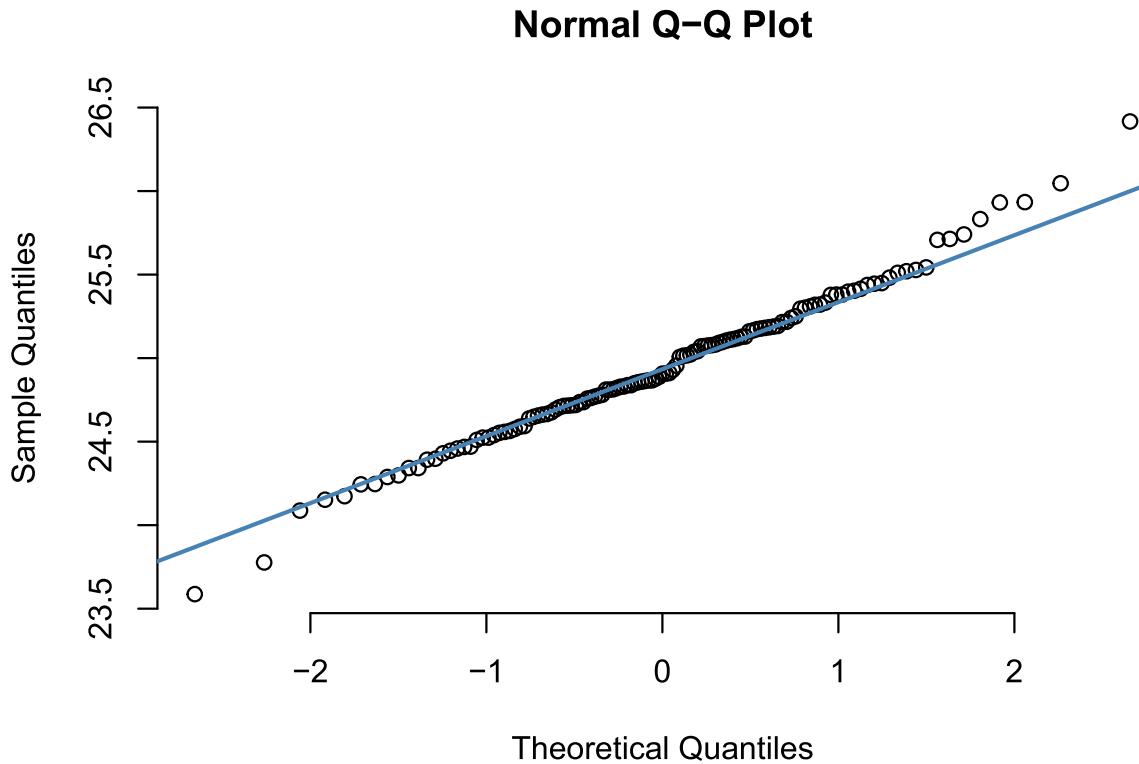


```
qqnorm(age1$Bilirubin, pch = 1, frame = FALSE)
qqline(age1$Bilirubin, col = "steelblue", lwd = 2)
```

Normal Q-Q Plot



```
qqnorm(age3$Bilirubin, pch = 1, frame = FALSE)
qqline(age3$Bilirubin, col = "steelblue", lwd = 2)
```



Since we see very straight lines for both residual plots, we can be ok with normal approximation. The qq-plots do show some deviance in the tails but since the data mostly hugs the line, we are still ok with normal approximation.

#(c) Let B_j be the expected difference in the log-concentration of bilirubin between recurrent wheezers and healthy controls at age j , and let $B_j\hat{}$ be its estimate you determined in part (a). Your collaborator has reason to believe that $B_1 = B_3 = B$. To estimate B , suppose you decide to meta-analyze the results at ages 1 and 3 by modeling $B_1\hat{}$, $B_3\hat{}$ as $\sim N(1B, \text{diag}(v_1\hat{}, v_3\hat{}))$, where $v_j\hat{}$ is the estimate for the variance of $B_j\hat{}$ determined in part (a). Assuming this model is correct, report a point estimate and 95% confidence interval for B .

```

pe = mean((mu1 - mu2), (mu1.2 - mu2.2))
t.2 = qtukey(p = 0.95, nmeans = 109, df = 18)/sqrt(2)
n1 = 109 + 55
n2 = 12 + 19
sigg = sig*sig.2
lo = pe - t*sigg
hi = pe + t*sigg
ci = c(lo, hi)
ci

## [1] 0.1108505 0.5606000

```

If we are assuming $B_1 = B_3$, and that the variance is the product of the variances of B_1 and B_3 estimates from (a), then a point estimate for B is 0.336 and a 95% CI is (0.111, 0.561).

#(d) Do you think the model in part (c) is appropriate? Why or why not? Given your answer, do you suspect the confidence interval determined in part (c) is too narrow, too wide, or accurate? Justify your answer.

The model is appropriate because of the assumptions we have covered, but the confidence interval may be too narrow because the variance will be estimated lower than the true variance, so we will have an anti-conservative confidence interval.

#(e) Let y_{ij} be the log-bilirubin concentration at age j in individual $i = 1, \dots, n_j$, and let x_{ji} be the covariates for individual i at age j used in part (a). Assume $E(y_{ij}) = x_{ji}T\gamma$...

(i) What do you expect the sign of ϕ to be?

If we are looking at the covariance of concentration for the same individual between two different ages, we would expect ϕ to be positive.

(ii) Based on the results from part (a), do you think it would be appropriate to let $\sigma_1^2 = \sigma_3^2$? Justify your answer.

Because the intervals from above do not overlap, we have no evidence to suggest they are different, so we can assume $\sigma_1^2 = \sigma_3^2$.

#(f) Using whatever method you deem most appropriate, estimate σ_j^2 and ϕ .

Using the same output from above, since we can assume equal variance between the ages, we have the following estimates for σ_j^2 and ϕ .

```
summary(fit1); summary(fit2)
```

```
##  
## Call:  
## lm(formula = Bilirubin ~ as.factor(Wheeze) + as.factor(Diet) +  
##       as.factor(Daycare) + as.factor(Sex), data = age1)  
##  
## Residuals:  
##      Min        1Q    Median        3Q       Max  
## -1.05025 -0.28395 -0.00135  0.30364  1.20615  
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 24.67417   0.12306 200.504 <2e-16 ***  
## as.factor(Wheeze)1 0.26449   0.17852   1.482   0.144  
## as.factor(Wheeze)2 -0.34211   0.27356  -1.251   0.216  
## as.factor(Diet)1  0.19134   0.22453   0.852   0.398  
## as.factor(Daycare)1 -0.18271   0.13270  -1.377   0.174  
## as.factor(Sex)1  0.06889   0.13259   0.520   0.605  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.5151 on 60 degrees of freedom  
## Multiple R-squared:  0.09042,    Adjusted R-squared:  0.01462  
## F-statistic: 1.193 on 5 and 60 DF,  p-value: 0.3235  
##  
## Call:  
## lm(formula = Bilirubin ~ as.factor(Wheeze) + as.factor(Diet) +  
##       as.factor(Daycare) + as.factor(Sex), data = age3)
```

```

## 
## Residuals:
##   Min     1Q Median     3Q    Max 
## -1.36463 -0.28971  0.01778  0.26659  1.46680
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)           25.00872   0.07628 327.844 <2e-16 ***
## as.factor(Wheeze)1  -0.08388   0.11918 -0.704   0.483    
## as.factor(Wheeze)2  -0.07911   0.16102 -0.491   0.624    
## as.factor(Diet)1    -0.04857   0.19283 -0.252   0.802    
## as.factor(Daycare)1 -0.05823   0.08245 -0.706   0.481    
## as.factor(Sex)1      -0.02253   0.08305 -0.271   0.787    
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.4584 on 121 degrees of freedom
## Multiple R-squared:  0.01309, Adjusted R-squared:  -0.02769 
## F-statistic: 0.3209 on 5 and 121 DF,  p-value: 0.8996

sig^2; sig.2^2

## [1] 0.07152178

## [1] 0.02545338

```

For age 1, $\sigma^2 = 0.072$ and for age 3, $\sigma^2 = 0.025$.

#(g) *See attached.*

#(h) *Using (e) and (f), estimate V. Comment on the off-diagonal elements of the matrix.*

If we estimate the variance from above of $B_{1\text{-hat}}$ and $B_{3\text{-hat}}$, the off-diagonal elements of this matrix will all be zero.

#(i) *Assume $B_{1\text{-hat}}$, $B_{3\text{-hat}}$ are jointly normal and that $B_1 = B_3 = B$. Use your estimate for V from part (h) to provide a point estimate and 95% CI for B.*

If $B_{1\text{-hat}}$ and $B_{3\text{-hat}}$ are jointly normal, their variance will be pooled and different from above, for a 95% CI.

3(g)

$$Y_j = (y_{j1}, \dots, y_{jn_j})^T$$

$$Z_{ij} = \begin{cases} 1, & \text{age } j \text{ recurrent wheeze} \\ 0, & \text{ow} \end{cases}$$

$$Z_j = (Z_{j1}, \dots, Z_{jn_j})^T$$

$$E(Y_j) = Z_j \beta_j + \tilde{X}_j \gamma_j$$

Show that for $\hat{\beta}_j$, OLS estimate for β_j

defined in (c) and $\tilde{Q}_j = I_{n_j} - \tilde{X}_j (\tilde{X}_j^T \tilde{X}_j)^{-1} \tilde{X}_j^T$

$$\hat{\beta}_j = (Z_j^T \tilde{Q}_j Z_j)^{-1} Z_j^T \tilde{Q}_j Y_j$$

$$= (Z_j^T [I_{n_j} - \tilde{X}_j (\tilde{X}_j^T \tilde{X}_j)^{-1} \tilde{X}_j^T] Z_j)^{-1} Z_j^T [I_{n_j} - \tilde{X}_j (\tilde{X}_j^T \tilde{X}_j)^{-1} \tilde{X}_j^T] Y_j$$

and from the earlier parts of this question we know that what Q will lead to is the OLS estimate for β .