**Seminar 3**

**Group A2**

**Yufeng Deng Yuanqing Wang**

**Task 1**

**Task 2**

**Task 3**

First, plot the change in log concentration over time by sex as a classification, observing its trend and linearity, as shown in the Figure3.1.

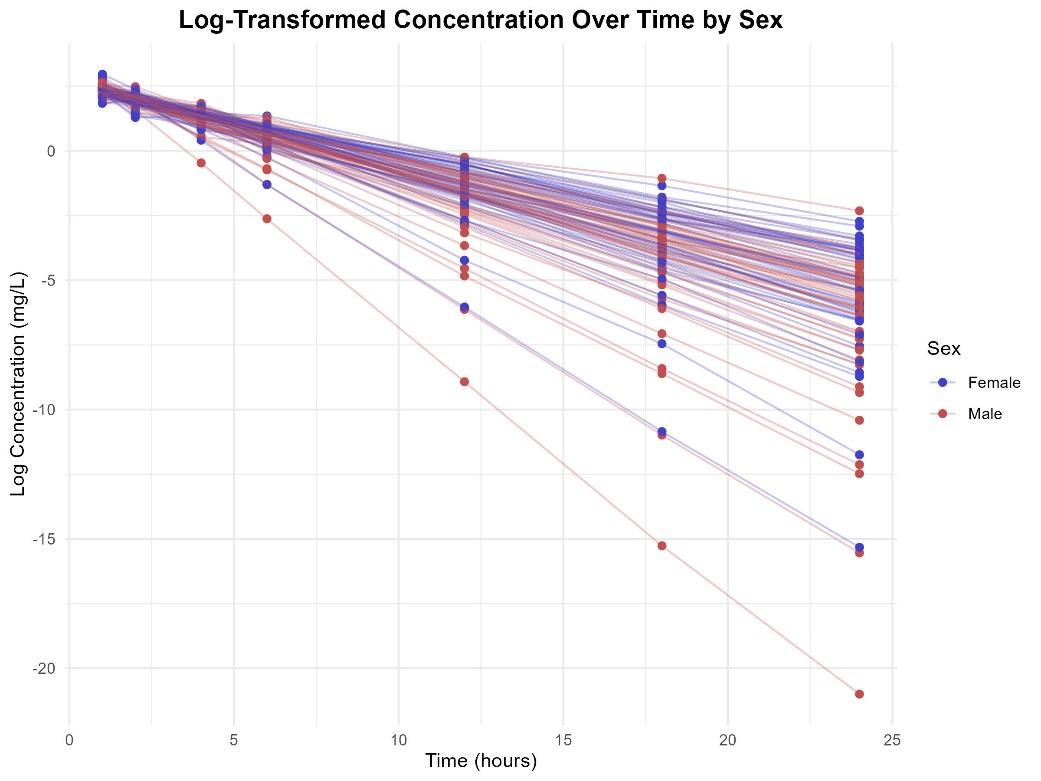


Figure3.1 Log Concentration Over Time

Through observation, it is evident that the linearity is significant. Furthermore, we use a linear mixed-effects model to analyze and evaluate the impact of various variables on the concentration-time change.

# Build the mixed-effects model with Time, BW, SEX, and AGE as fixed effects, and ID as a random effect

model <- lmer(log\_DV ~ Time + Bw + Sex + Age + (1 | ID), data = data)

# Display model summary to examine the significance of each fixed effect

summary(model)

This code constructs a linear mixed-effects model that incorporates both fixed and random effects, where Time, BW (body weight), SEX (gender), and AGE (age) are treated as fixed effects to assess their impact on log concentration (log\_DV). Additionally, ID (each patient) is included as a random effect, allowing for individual variability in concentration changes and enhancing the model’s fit by accounting for patient-specific deviations in log concentration trends.

Random effects:

Groups Name Variance Std.Dev.

ID (Intercept) 0.7519 0.8671

Residual 1.2377 1.1125

Number of obs: 700, groups: ID, 100

Fixed effects:

Estimate Std. Error t value

(Intercept) 7.986750 2.290779 3.486

Time -0.364388 0.005189 -70.227

Bw -0.043189 0.022434 -1.925

Sex -0.111348 0.218560 -0.509

Age -0.027994 0.024580 -1.139

Correlation of Fixed Effects:

(Intr) Time Bw Sex

Time -0.022

Bw -0.771 0.000

Sex 0.346 0.000 -0.460

Age -0.618 0.000 -0.019 -0.041

The results indicate that time has a strong and highly significant negative effect on log concentration, suggesting that as time progresses, log concentration decreases significantly. Body weight also shows a negative relationship with log concentration, although the effect is weaker and only marginally significant. In contrast, sex and age do not exhibit significant effects on log concentration, with the estimates being small and the corresponding t-values not reaching statistical significance. The random effects analysis reveals that there is significant individual variability in baseline log concentration, as indicated by the substantial variance in the random intercept for patients. The residual variance reflects within-patient variability in log concentration changes over time. These findings highlight the importance of time and body weight in influencing log concentration, while suggesting that sex and age may not be as relevant in this context.

**Task 4**

For this task, we used the built-in R libraries **survival** and **survminer**.

library(survival)

library(survminer)

The survival package in R is a fundamental tool for survival analysis, providing functions for the estimation and modeling of survival data. It supports Kaplan-Meier estimation, Cox proportional hazards models, and parametric survival models (e.g., Weibull). The package is widely used for analyzing time-to-event data, particularly in clinical research and epidemiology. Key functions include Surv() for creating survival objects, survfit() for estimating survival curves, and coxph() for fitting Cox regression models.

The survminer package is an extension designed for visualizing survival analysis results. It integrates well with the survival package and offers various functions for creating high-quality survival plots, such as Kaplan-Meier curves and Cox model diagnostics. The ggsurvplot() function is a key tool for visualizing survival curves, and other functions help assess the proportional hazards assumption and generate diagnostic plots for survival models.

data <- read.csv("data\_task4.csv")

surv\_obj <- Surv(time = data$Time, event = data$Status)

surv\_diff <- survdiff(surv\_obj ~ Group, data = data)

print(surv\_diff)

cox\_model <- coxph(surv\_obj ~ Group, data = data)

summary(cox\_model)

fit <- survfit(surv\_obj ~ Group, data = data)

ggsurvplot(fit, data = data, pval = TRUE,

           xlab = "Time (Months)", ylab = "Survival Probability",

           legend.labs = c("Control", "Treatment"),

           legend.title = "Group")

Then, we obtain the following results.

Call:

survdiff(formula = surv\_obj ~ Group, data = data)

N Observed Expected (O-E)^2/E (O-E)^2/V

Group=Control 100 73 90.6 3.43 9.53

Group=Treatment 100 72 54.4 5.72 9.53

Chisq= 9.5 on 1 degrees of freedom, p= 0.002

Call:

coxph(formula = surv\_obj ~ Group, data = data)

n= 200, number of events= 145

coef exp(coef) se(coef) z Pr(>|z|)

GroupTreatment 0.5204 1.6828 0.1703 3.055 0.00225 \*\*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

exp(coef) exp(-coef) lower .95 upper .95

GroupTreatment 1.683 0.5943 1.205 2.35

Concordance= 0.573 (se = 0.023 )

Likelihood ratio test= 9.26 on 1 df, p=0.002

Wald test = 9.34 on 1 df, p=0.002

Score (logrank) test = 9.53 on 1 df, p=0.002

The results of the survival analysis reveal a statistically significant difference in survival between the Control and Treatment groups. The log-rank test (Chi-square = 9.5, p = 0.002) indicates that the two groups have significantly different survival profiles.

Further analysis with the Cox proportional hazards model shows that the hazard ratio (HR) for the Treatment group is 1.683 (95% CI: 1.205 - 2.350), suggesting that the Treatment group has a 68.3% higher risk of experiencing the event (death) compared to the Control group. This result is statistically significant (p = 0.00225). The concordance index of 0.573 suggests that the model has moderate predictive ability.

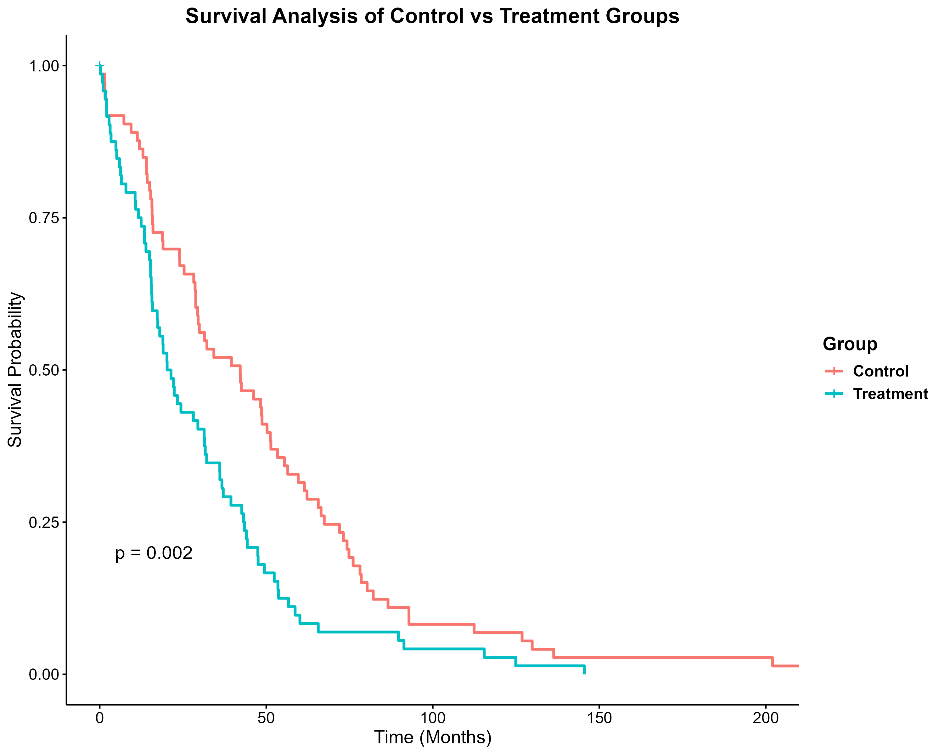


Figure 4.1 Survival Analysis of Control and Treatment Group

Despite the significant survival difference observed in the log-rank test, the Cox model indicates that the Treatment group has a higher risk of death. This suggests that the experimental treatment may not improve overall survival compared to the control. As shown in the Figure 4.1, the Control group has better survival rates at the same time points, with a significantly lower risk of death. In contrast, the Treatment group exhibits a higher risk of death. This result suggests that the experimental treatment may not have significantly improved patient survival, and it may even be worse than the Control group.