STA 3180 Statistical Modelling: Longitudinal Data

Longitudinal Data in STA 3180 Statistical Modelling

Longitudinal data is a type of data that is collected over time, with multiple observations taken from the same individuals or entities. It is a powerful tool for analyzing changes in behavior, attitudes, and other characteristics over time. In this lecture, we will discuss the basics of longitudinal data, including its structure, analysis techniques, and coding examples.

What is Longitudinal Data?

Longitudinal data is a type of data that is collected over time, with multiple observations taken from the same individuals or entities. It can be used to track changes in behavior, attitudes, and other characteristics over time. It is often collected in surveys, experiments, or observational studies.

Types of Longitudinal Data

There are two main types of longitudinal data: panel data and repeated measures data.

Panel Data: Panel data is a type of longitudinal data where the same group of individuals or entities is observed over a period of time. This type of data is useful for tracking changes in behavior, attitudes, and other characteristics over time.

Repeated Measures Data: Repeated measures data is a type of longitudinal data where the same individuals or entities are observed multiple times over a period of time. This type of data is useful for studying changes in behavior, attitudes, and other characteristics over time.

Structure of Longitudinal Data

Longitudinal data is usually structured as a data frame, with each row representing an individual or entity, and each column representing an observation. The columns may include variables such as time, age, gender, and other characteristics.

Analysis Techniques

There are several techniques for analyzing longitudinal data, including linear regression, mixed-effects models, and survival analysis.

Linear Regression: Linear regression is a statistical technique used to analyze the relationship between one or more independent variables and a dependent variable. It can be used to analyze changes in a dependent variable over time.

Mixed-Effects Models: Mixed-effects models are a type of regression model that includes both fixed and random effects. They are useful for analyzing changes in a dependent variable over time.

Survival Analysis: Survival analysis is a type of statistical analysis used to analyze the time until an event occurs. It can be used to analyze changes in behavior, attitudes, and other characteristics over time.

Coding Examples

```
Start of Code
# Load libraries
library(tidyverse)
library(lme4)
# Read in data
data <- read_csv("longitudinal_data.csv")
# Fit linear regression model
model <- lm(Dependent_Variable ~ Time, data = data)
# Fit mixed-effects model
model <- lmer(Dependent_Variable ~ Time + (1 | Entity), data = data)
# Fit survival analysis model
model <- survreg(Surv(Time, Event) ~ Age + Gender, data = data)
End of Code
```

Practice Multiple Choice Questions

Q1. What type of data is longitudinal data?

A. Panel data

B. Time series data

C. Cross-sectional data

D. Repeated measures data

Answer: A. Panel data