

Uniwersytet Wrocławski  
Wydział Matematyki i Informatyki  
Instytut Matematyczny  
*specjalność: Analiza Danych*

*Magda Kozajda*

**Statystyczna analiza danych  
z użyciem pakietu nparLD**

Praca licencjacka  
napisana pod kierunkiem  
Pana dra Grzegorza Wyłupka

Wrocław 2023

Składam serdeczne podziękowania  
Panu doktorowi Grzegorzowi Wyłupkowi  
za wszystkie wskazówki,  
za ciągłą możliwość konsultacji  
i po prostu - za zainteresowanie mnie statystyką.

# Spis treści

<b>1</b>	<b>Wstęp</b>	<b>5</b>
1.1	Motywacja . . . . .	5
<b>2</b>	<b>Modele nieparametryczne</b>	<b>7</b>
2.1	Ogólny model nieparametryczny . . . . .	7
2.2	Wybrane modele nieparametryczne . . . . .	7
2.2.1	LD-F1 . . . . .	8
2.2.2	F1-LD-F1 . . . . .	8
2.2.3	F2-LD-F1 . . . . .	9
<b>3</b>	<b>Statystyki i testowanie</b>	<b>11</b>
3.1	Model i motywacja . . . . .	11
3.2	Estymacja rangowa . . . . .	12
3.2.1	Statystyka $p_s$ . . . . .	12
3.2.2	Estymacja $p_s$ . . . . .	13
3.3	Testowanie hipotez . . . . .	14
3.3.1	Statystyka typu Walda . . . . .	14
3.3.2	Statystyka typu ANOVA . . . . .	15
<b>4</b>	<b>Przykładowy eksperyment</b>	<b>16</b>
4.1	Model . . . . .	16
4.2	Testowanie . . . . .	17
4.2.1	Efekt grupy . . . . .	17
4.2.2	Efekt czasu . . . . .	18
4.2.3	Interakcja . . . . .	18
<b>5</b>	<b>Przykłady</b>	<b>19</b>
5.1	Badanie ze stomatologii . . . . .	19
5.1.1	Analiza nieparametryczna . . . . .	19
5.1.2	Analiza parametryczna . . . . .	22
5.2	Badanie rozwoju szczura . . . . .	24
5.3	Badanie zaburzeń oddychania . . . . .	27
	<b>Spis tabel</b>	<b>32</b>
	<b>Spis wykresów</b>	<b>33</b>
	<b>Literatura</b>	<b>34</b>

## Abstrakt

Eksperymenty naukowe często wymagają wielokrotnego egzaminowania tych samych obiektów przez długi okres czasu. Zebrane w ten sposób dane longitudinalne są zależne, mają obserwacje odstające i często są niekompletne. Dodatkową trudnością w ich analizie jest, że nie znamy rozkładu, z którego pochodzą i nie możemy estymować brakujących parametrów. W związku z tym do testowania statystycznego będziemy używać jedynie wiedzy o liczbie obserwacji i o ich dystrybuantach. Procedura pozwalająca na wyliczenie statystyk (typu Walda i typu ANOVA) oparta jest na idei modelu nieparametrycznego i na metodzie rangowania. Pakiet `nparLD` w środowisku R pozwala na praktyczną obsługę danych longitudinalnych. W niniejszej pracy opiszemy teoretyczne zagadnienia, na których oparty jest pakiet `nparLD` i zaprezentujemy jego działanie na przykładach z nauk przyrodniczych.

**Słowa kluczowe:** dane longitudinalne, model nieparametryczny, efekt grupy, efekt czasu, interakcja, `nparLD`, R.

# Rozdział 1

## Wstęp

### 1.1 Motywacja

W naukach biologicznych, medycznych czy socjologicznych często mamy do czynienia z eksperymentami, w których obiekt badany jest wielokrotnie, w różnych punktach czasu. Dane wymagające częstego egzaminowania tych samych podmiotów, w ten sam sposób i w określonych odstępach czasu, nazywamy danymi longitudinalnymi. Pomiar z tej samej jednostki doświadczalnej są zależne, co znacznie komplikuje ich statystyczną analizę. W pracy z 1994 roku P. Diggle, K. Liang i S. Zeger obszernie opisali metody badań dla danych longitudinalnych. Ich podejście oparte jest głównie na uogólnionych modelach liniowych (GLM) i ich rozszerzeniach (GLMM). Implementacje tych procedur są zazwyczaj ogólnodostępne w oprogramowaniu typu SAS czy R. Korzystanie z istniejących tam narzędzi może być problematyczne, ponieważ często oparte są one na dodatkowych założeniach, które nie muszą być spełnione w praktyce. Z tego powodu metody parametryczne mogą prowadzić do błędnych wniosków.

Innym podejściem do analizy danych longitudinalnych są nieparametryczne metody rangowe. Do ich badania, tworzenia i rozwoju znacząco przyczyniają się: E. Brunner, F. Konietzschke, A. Bathke i M. Puri. W swoich publikacjach podkreślają oni liczne zalety metod nieparametrycznych. Są to między innymi: odporność na obserwacje odstające, efektywność dla prób o małych rozmiarach czy ogólna elastyczność w stosowaniu dla danych różnego rodzaju. W pracy z 2002 roku E. Brunner szczegółowo opisał różnorodne nieparametryczne metody rangowe dla danych longitudinalnych. Przybliżone są tam również metody estymowania statystyk i testowania hipotez oraz wyznaczania przedziałów ufności.

Dane longitudinalne często gromadzone są w eksperymentach czynnikowych, czyli takich, których budowa składa się z dwóch lub więcej czynników. Przykładowe eksperymenty tego typu to:

- Jedna homogeniczna grupa obiektów obserwowana wielokrotnie w  $s = 1, \dots, t$  punktach czasu.
- Jedna homogeniczna grupa obiektów obserwowana wielokrotnie w  $s = 1, \dots, t$  punktach czasu, gdzie pomiary dla każdego obiektu z grupy zbierane są osobno. Może być to przykładowo badanie oczu, gdzie dla każdego pomiaru zapisujemy wynik dla lewego i dla prawego oka.
- Jedna homogeniczna grupa obiektów obserwowana wielokrotnie pod  $l = 1, \dots, c$  warunkami, gdzie każdy obiekt z grupy obserwowany jest pod każdym z warunków w  $s = 1, \dots, t$  punktach czasu.

- Więcej niż jedna homogeniczna grupa obiektów obserwowana w  $s = 1, \dots, t$  punktach czasu. Przykładem jest podział badania na grupę kobiet i mężczyzn lub grupę kontrolną i zabiegową.

Analiza danych longitudinalnych opiera się głównie na analizie ich dystrybuant. Dystrybuanty są także niezbędne do wnioskowania o wynikach eksperymentu oraz do testowania hipotez statystycznych. Typowymi problemami, które chcemy badać są:

- Czy różne metody leczenia dają ten sam efekt?
- Czy upływ czasu ma wpływ na poprawę lub pogorszenie się wyników leczenia?
- Czy wyniki leczenia dla wszystkich grup są identyczne przez cały okres badania?

Te pytania można łatwo przetłumaczyć na język statystyczny. Testowanie hipotez dla tego typu danych skupia się głównie na: badaniu efektu czasu, efektu podziału na grupy oraz na badaniu interakcji między czasem a różnymi grupami. M. Akritas i G. Arnold w swojej pracy z 1997 roku jako jedni z pierwszych sformułowali te hipotezy oraz sposób ich testowania.

Jednym z narzędzi do analizy danych longitudinalnych jest pakiet `nparLD` dostępny w oprogramowaniu R. Jego nazwa złożona jest z dwóch członów: „npar” - pochodzącego od słowa „nonparametric” (nieparametryczne) oraz „LD” będącego skrótem od „longitudinal data” (dane longitudinalne).

Niniejsza praca podzielona jest na dwie części: teoretyczną i praktyczną. Część teoretyczna zaprezentowana jest w rozdziałach: drugim, trzecim i czwartym. W rozdziale 2 opisane zostały wybrane typy modeli dla danych nieparametrycznych. Rozdział 3 zawiera testowanie hipotez oraz teoretyczne wyliczenia statystyk typu Walda i typu ANOVA. W rozdziale 4 za pomocą przykładowego eksperymentu pokazana jest intuicja dotycząca testowania. Rozdział piąty jest częścią praktyczną pracy. Zawiera opis pakietu `nparLD` oraz przykładowe eksperymenty wykonane przy jego użyciu.

## Rozdział 2

# Modele nieparametryczne

W tym rozdziale opiszemy modele nieparametryczne stosowane przy analizie danych longitudinalnych. Na początku poznamy ideę ogólnego modelu nieparametrycznego. Następnie omówimy najczęściej używane modele i skonstruujemy dla nich schematy. Wprowadzimy także system notacyjny pozwalający na szybki opis budowy modelu przez uwzględnienie liczby występujących w nim czynników.

### 2.1 Ogólny model nieparametryczny

Modele parametryczne zakładają istnienie rozkładu opisywalnego różnorodnymi parametrami. W modelach nieparametrycznych nie mamy żadnych szczegółowych informacji na temat rozkładu. Do estymacji jego parametrów musimy bezpośrednio używać danych.

Założmy, że badamy  $n$  obiektów w  $t$  punktach czasu. Wyniki zebrane od tej samej,  $k$ -tej jednostki,  $k = 1, \dots, n$ , przez cały okres badania są zależne. Wyniki zebrane od  $n$  różnych jednostek w chwili  $s$  są niezależne o dystrybuancie  $F_s$  dla  $s = 1, \dots, t$ . W modelu nieparametrycznym do testowania statystycznego używamy jedynie wiedzy o liczbie obserwacji i o dystrybuantach  $F_s$ .

Poniższa definicja w ściślejszy sposób przedstawia ideę ogólnego modelu nieparametrycznego:

#### Definicja 2.1

Niech  $X_k = (X_{k1}, \dots, X_{kt})'$ ,  $k = 1, \dots, n$  będą niezależnymi wektorami losowymi, a  $F_s$  dystrybuantą brzegową, taką że:  $X_{ks} \sim F_s$ ,  $k = 1, \dots, n$ ,  $s = 1, \dots, t$ . Łączny wektor dystrybuant brzegowych w  $t$  punktach czasu to  $F = (F_1, \dots, F_t)'$ .

Model nazywamy **modelem nieparametrycznym** wtedy i tylko wtedy, gdy do formułowania hipotez używamy jedynie dystrybuant brzegowych, czyli wektora  $F$ .

### 2.2 Wybrane modele nieparametryczne

Ogólny model nieparametryczny należy każdorazowo dopasować do przeprowadzanego eksperymentu. W celu opisanego najczęściej używanych modeli wprowadzimy system notacyjny  $F_x$ -LD- $F_y$ . Skrót „LD” jest tutaj rozwinięciem określenia „Dane Longitudinalne”. Symbol  $x$  oznacza liczbę czynników whole-plot, czyli takich, które dzielą próbę na niezależne grupy. Symbol  $y$  to liczba czynników sub-plot, czyli czynników dzielących próbę na grupy powtarzających się pomiarów. Przykładowo: dla grup kobiet i mężczyzn badanych trzykrotnie w ciągu dnia, płeć będzie czynnikiem whole-plot, a czas czynnikiem sub-plot. Przykładowe schematy (LD-F1, F1-LD-F1, F2-LD-F1) i ich opisy można znaleźć poniżej.

### 2.2.1 LD-F1

Rozważamy eksperyment, w którym grupa  $n$  obiektów jest badana  $t$  razy. Przykładem może być próba kliniczna Bandelowa z 1998 roku. Polegała ona na mierzeniu wskaźnika CGI (wskaźnik efektywności leczenia u osób chorych psychicznie) u 16 pacjentów przez 8 tygodni. Tabela 2.1 przedstawia schemat dla obserwacji  $X_{ks}$  i ich rozkładów brzegowych  $F_s$ .

	Dane			Rozkłady brzegowe		
	Czas			Czas		
Obserwacja	$s = 1$	...	$s = t$	$s = 1$	...	$s = t$
$k = 1$	$X_{11}$	...	$X_{1t}$	$F_1$	...	$F_t$
$\vdots$	$\vdots$			$\vdots$		
$k = n$	$X_{n1}$	...	$X_{nt}$	$F_1$	...	$F_t$

Tabela 2.1: Model LD-F1

Dane - wektor losowy:

$$X_k = (X_{k1}, \dots, X_{kt})', \quad k = 1, \dots, n.$$

Rozkład brzegowy:

$$X_{1s}, \dots, X_{ns} \sim F_s, \quad s = 1, \dots, t,$$

gdzie  $F_s$  oznacza dystrybuantę brzegową z próby  $s$ .

W modelu mamy  $n$  wektorów, w każdym z nich zbieramy  $t$  danych, co daje łącznie

$$N = n \cdot t$$

zmiennych.

Dla modelu LD-F1 możemy testować hipotezę o „braku efektu czasu” przedstawioną w języku dystrybuant jako:

$$H_0^T : F_1 = \dots = F_t.$$

### 2.2.2 F1-LD-F1

W tej sytuacji  $a$  różnych grup jest obserwowanych w  $t$  punktach czasu. Przykładem jest badanie Sedghiego o efektywności technik irygacyjnych przy usuwaniu miazgi zębowej podczas leczenia kanałowego (2010). W eksperymencie 30 osób podzielono losowo na trzy równe grupy. Każda z nich została poddana leczeniu z inną procedurą oczyszczania kanałów zębowych. Następnie zmierzono pozostały poziom miazgi i wybrano najbardziej efektywną technikę. W tabeli 2.2 przedstawiono schemat dla  $X_{iks}$  obserwacji i ich dystrybuant brzegowych  $F_{is}$ .

Dane - niezależne wektory losowe:

$$X_{ik} = (X_{ik1}, \dots, X_{ikt})', \quad k = 1, \dots, n_i.$$

Rozkład brzegowy:

$$X_{iks} \sim F_{is}, \quad i = 1, \dots, a, \quad s = 1, \dots, t.$$



		Dane			Rozkład		
		Czas			Czas		
Czynnik A	Obserwacja	$s = 1$	...	$s = t$	$s = 1$	...	$s = t$
$i = 1$	$k = 1$	$X_{111}$	...	$X_{11t}$	$F_{11}$	...	$F_{1t}$
	$\vdots$		$\vdots$			$\vdots$	
	$k = n_1$	$X_{1n_11}$	...	$X_{1n_1t}$	$F_{11}$	...	$F_{1t}$
$\vdots$	$\vdots$		$\vdots$			$\vdots$	
$i = a$	$k = 1$	$X_{a11}$	...	$X_{a1t}$	$F_{a1}$	...	$F_{at}$
	$\vdots$		$\vdots$			$\vdots$	
	$k = n_a$	$X_{an_a1}$	...	$X_{an_at}$	$F_{a1}$	...	$F_{at}$

Tabela 2.2: Model F1-LD-F1

W modelu mamy  $n = \sum_{i=1}^a n_i$  wektorów, w każdym z nich zbieramy  $t$  danych, co daje łącznie:

$$N = n \cdot t = \sum_{i=1}^a n_i \cdot t$$

zmiennych.

Dla modelu F1-LD-F1 możemy testować hipotezy o: braku efektu grupy (A), braku efektu czasu (T) i braku interakcji (AT) między A i T. Hipotezy przedstawione są w języku dystrybuant jako:

$$H_0^A : \overline{F}_{1\cdot} = \dots = \overline{F}_{a\cdot},$$

$$H_0^T : \overline{F}_{\cdot 1} = \dots = \overline{F}_{\cdot t},$$

$$H_0^{AT} : F_{is} = \overline{F}_{i\cdot} - \overline{F}_{\cdot s} + \overline{F}_{\cdot\cdot}, \quad i = 1, \dots, a, \quad s = 1, \dots, t,$$

gdzie  $\overline{F}_{i\cdot} = \frac{1}{t} \sum_{s=1}^t F_{is}$  jest uśrednioną dystrybuantą po czasie dla grupy  $i$ ,  $\overline{F}_{\cdot s} = \frac{1}{a} \sum_{i=1}^a F_{is}$  to uśredniona dystrybuanta po grupach  $i$  w punkcie czasu  $s$  oraz  $\overline{F}_{\cdot\cdot} = \frac{1}{at} \sum_{i=1}^a \sum_{s=1}^t F_{is}$  to uśredniona dystrybuanta po wszystkich parametrach.

### 2.2.3 F2-LD-F1

W modelu F2-LD-F1 obserwujemy  $a \cdot b$  różnych grup w  $t$  punktach czasu. Przykładem może być eksperyment, w którym najpierw dzielimy uczestników według płci a następnie każdą z podgrup dzielimy według losowo dobranej metody leczenia. Podczas trwania eksperymentu wielokrotnie zbieramy dane o wynikach kuracji w różnych punktach czasu  $t$ . W tabeli (2.3) zobrazowany został schemat dla obserwacji  $X_{ijks}$  i dla odpowiadających im dystrybuant brzegowych  $F_{ijs}$ .

Dane - niezależne wektory losowe:

$$X_{ijk} = (X_{ijk1}, \dots, X_{ijk t})', \quad k = 1, \dots, n_{ij}.$$

Rozkład brzegowy:

$$X_{ijks} \sim F_{ijs}, \quad i = 1, \dots, a, \quad j = 1, \dots, b, \quad s = 1, \dots, t.$$

			Dane			Rozkład		
			Czas			Czas		
Czynnik A	Czynnik B	Obserwacja	$s = 1$	...	$s = t$	$s = 1$	...	$s = t$
$i = 1$	$j = 1$	$k = 1$	$X_{1111}$	...	$X_{111t}$	$F_{111}$	...	$F_{11t}$
		$\vdots$			$\vdots$		$\vdots$	
	$j = b$	$k = n_{11}$	$X_{11n_{11}1}$	...	$X_{11n_{11}t}$	$F_{111}$	...	$F_{11t}$
		$\vdots$	$\vdots$		$\vdots$		$\vdots$	
		$k = 1$	$X_{1b11}$	...	$X_{1b1t}$	$F_{1b1}$	...	$F_{1bt}$
		$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$j = b$	$k = n_{1b}$	$X_{1bn_{1b}1}$	...	$X_{1bn_{1b}t}$	$F_{1b1}$	...	$F_{1bt}$	
	$\vdots$	$\vdots$		$\vdots$		$\vdots$		
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		
$i = a$	$j = 1$	$k = 1$	$X_{a111}$	...	$X_{a11t}$	$F_{a11}$	...	$F_{a1t}$
		$\vdots$			$\vdots$		$\vdots$	
	$j = b$	$k = n_{a1}$	$X_{a1n_{a1}1}$	...	$X_{a1n_{a1}t}$	$F_{a11}$	...	$F_{a1t}$
		$\vdots$	$\vdots$		$\vdots$		$\vdots$	
		$k = 1$	$X_{ab11}$	...	$X_{ab1t}$	$F_{ab1}$	...	$F_{abt}$
		$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$j = b$	$k = n_{ab}$	$X_{abn_{ab}1}$	...	$X_{abn_{ab}t}$	$F_{ab1}$	...	$F_{abt}$	
	$\vdots$	$\vdots$		$\vdots$		$\vdots$		

Tabela 2.3: Model F2-LD-F1

W modelu mamy  $n = \sum_{i=1}^a \sum_{j=1}^b n_{ij}$  obserwacji, w każdej z nich zbieramy  $t$  danych, co daje łącznie

$$N = n \cdot t = \sum_{i=1}^a \sum_{j=1}^b n_{ij} \cdot t$$

zmiennych.

Dla modelu F2-LD-F1 możemy testować hipotezy o:

- braku efektu grupy (A) i braku efektu grupy (B),
- braku efektu czasu (T),
- braku interakcji (AT) między A i T, braku interakcji (BT) między B i T, braku interakcji (AB) między A i B oraz o braku interakcji (ABT) między A, B i T.

Hipotezy stawiane są analogicznie jak w modelu F1-LD-F1.

## Rozdział 3

# Statystyki i testowanie

W rozdziale trzecim omówimy jeden ze sposobów wyliczania statystyk dla danych longitudinalnych, czyli estymację z użyciem rang. Przedstawimy także procedury testowania hipotez dla danych nieparametrycznych.

### 3.1 Model i motywacja

Na początku przypomnijmy jak wygląda ogólny model nieparametryczny typu LD-F1:

	Dane			Rozkład		
	Czas			Czas		
Obserwacja	$s = 1$	...	$s = t$	$s = 1$	...	$s = t$
$k = 1$	$X_{11}$	...	$X_{1t}$	$F_1$	...	$F_t$
$\vdots$		$\vdots$			$\vdots$	
$k = n$	$X_{n1}$	...	$X_{nt}$	$F_1$	...	$F_t$

Tabela 3.1: Model LD-F1

W eksperymencie odpowiadającym temu modelowi badamy  $k = 1, \dots, n$  obiektów, każdy  $t$ -krotnie, co daje łącznie  $N = n \cdot t$  obserwacji. Dane zebrane dla  $k$ -tego obiektu przedstawione są jako wektor:

$$X_k = (X_{k1}, \dots, X_{kt})', \quad k = 1, \dots, n.$$

Rozkład brzegowy dla tych danych to:

$$X_{1s}, \dots, X_{ns} \sim F_s, \quad s = 1, \dots, t,$$

gdzie  $F_s$  oznacza dystrybuantę brzegową z próby  $s = 1, \dots, t$ .

Widzimy, że model nie zawiera żadnych parametrów, które opisywałyby różnicę pomiędzy dystrybuantami brzegowymi:  $F_1, \dots, F_t$ . Nie mamy zatem parametrów niezbędnych do formułowania i testowania hipotez a zarazem do analizy wyników eksperymentu. Aby rozwiązać ten problem wprowadzamy estymację z użyciem rang.

## 3.2 Estymacja rangowa

Zanim przejdziemy do idei estymacji rangowej, przypomnimy potrzebne definicje i wykonamy odpowiednie przekształcenia.

Rozważamy zbiór niezależnych zmiennych losowych  $X_{ks} \sim F_s(x)$  dla  $k = 1, \dots, n$  i  $s = 1, \dots, t$ . Przypomnijmy, że  $F_s(x)$  jest dystrybuantą, czyli funkcją jednoznacznie określającą rozkład prawdopodobieństwa danych  $X_{ks}$ ,  $k = 1, \dots, n$ , w chwili  $s$ ,  $s = 1, \dots, t$ . Na potrzeby pracy dystrybuantę definiujemy jako:

$$F(x) = \frac{1}{2}[F^+(x) + F^-(x)], \quad (3.1)$$

gdzie  $F^+(x) = P(X \leq x)$  jest prawostronnie ciągłą wersją dystrybuanty, a  $F^-(x) = P(X < x)$  wersją lewostronnie ciągłą.

Stąd dla  $s = 1, \dots, t$  mamy:

$$F_s(x) = \frac{1}{2}[P(X_{1s} \leq x) + P(X_{1s} < x)] = P(X_{1s} < x) + 0.5P(X_{1s} = x). \quad (3.2)$$

Wprowadźmy teraz funkcję  $H(x) = \frac{1}{t} \sum_{s=1}^t F_s(x)$  będącą średnią wszystkich dystrybuant występujących w eksperymencie. Korzystając ze wzoru (3.2) możemy zapisać  $H(x)$  jako:

$$H(x) = \frac{1}{t} \sum_{s=1}^t F_s(x) = \frac{1}{t} \sum_{s=1}^t P(X_{1s} < x) + 0.5P(X_{1s} = x). \quad (3.3)$$

Dla danych pochodzących z rozkładu o dystrybuancie  $F_s(x)$ ,  $s = 1, \dots, t$  i średniej dystrybuant równej  $H(x)$  wprowadzamy statystykę  $p_s$  nazywaną względnym efektem leczenia (relative marginal effects).

### 3.2.1 Statystyka $p_s$

Chcemy opisać tendencję w zachowaniu  $F_s(x)$ ,  $s = 1, \dots, t$  względem  $H(x)$ . W tym celu całkujemy funkcję  $H(x)$  względem  $F_s$ :

$$p_s = \int_{\mathbb{R}} H(x) dF_s(x). \quad (3.4)$$

Korzystając z wyżej wyprowadzonych  $F_s(x) = P(X_{1s} < x) + 0.5P(X_{1s} = x)$  i  $H(x) = \frac{1}{t} \sum_{s=1}^t P(X_{1s} < x) + 0.5P(X_{1s} = x)$  możemy przekształcić równanie (3.4) do postaci:

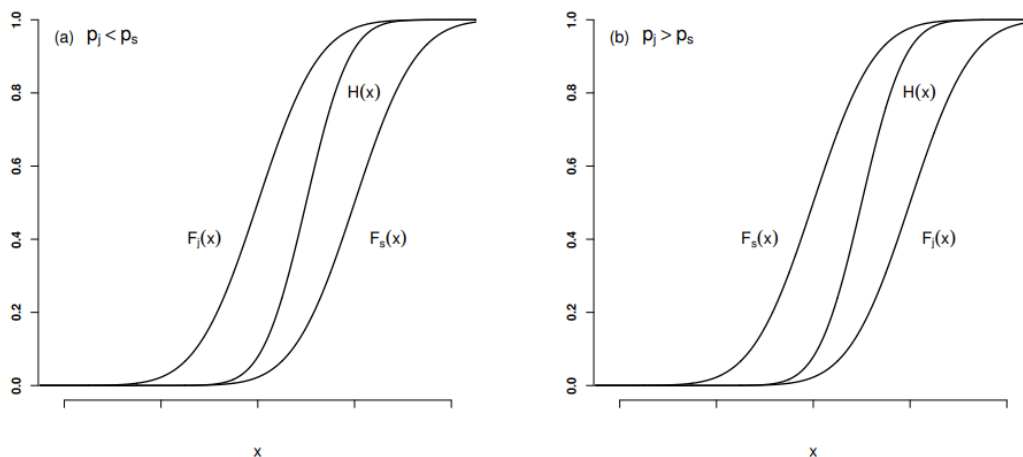
$$p_s = \int_{\mathbb{R}} H(x) dF_s(x) = P(Z < X_{1s}) + 0.5P(Z = X_{1s}), \quad s = 1, \dots, t, \quad (3.5)$$

gdzie  $Z \sim H$  oznacza losowo wybraną obserwację z całego zbioru danych niezależnie od  $X_{1s}$ . Statystykę  $p_s$  można interpretować jako prawdopodobieństwo, że losowo wybrana obserwacja  $X_{ks}$  w punkcie czasu  $s$  ma tendencję do bycia większą niż  $Z$ .

Dla dowolnych obserwacji  $X_{1j}$  i  $X_{2s}$  zachodzi także:

- jeżeli  $p_j < p_s$ , to  $X_{1j} < X_{2s}$  z dużym prawdopodobieństwem,
- jeżeli  $p_j > p_s$ , to  $X_{1j} > X_{2s}$  z dużym prawdopodobieństwem,
- jeżeli  $p_j = p_s$ , to  $X_{1j} = X_{2s}$  z dużym prawdopodobieństwem.

Na rysunku (3.1) przedstawione są dystrybuanty  $F_s(x)$ ,  $F_j(x)$  oraz średnia dystrybuanta  $H(x)$ . Grafika daje intuicję na temat związku dystrybuant  $F_s(x)$ ,  $F_j(x)$  ze zmiennymi  $p_j$  i  $p_s$ .

Rysunek 3.1: Związek statystyki  $p$  z dystrybuantą  $F$ .

Źródło: Journal of Statistical Software (2012).

### 3.2.2 Estymacja $p_s$

W celu estymacji  $p_s$  dokonujemy rangowania, czyli zastępujemy wszystkie obserwacje  $X_{11}, \dots, X_{nt}$  przez ich rangi  $R_{11}, \dots, R_{nt}$ . Wprowadźmy zmienną  $\bar{R}_{.s} = \frac{1}{n} \sum_{k=1}^n R_{ks}$ , oznaczającą średnią rang w próbie  $s$ .

	Dane			Średnia Rang		
	Czas			Czas		
Obserwacja	$s = 1$	...	$s = t$	$s = 1$	...	$s = t$
$k = 1$	$R_{11}$	...	$R_{1t}$	$\bar{R}_{.1}$	...	$\bar{R}_{.t}$
$\vdots$	$\vdots$					
$k = n$	$R_{n1}$	...	$R_{nt}$			

Tabela 3.2: Rangowanie danych z modelu LD-F1

Do estymacji dystrybuanty  $F_s$ ,  $s = 1, \dots, t$  użyjemy jej empirycznego odpowiednika:

$$\hat{F}_s(x) = \frac{1}{2}[\hat{F}^+(x) + \hat{F}^-(x)] = \frac{1}{n} \sum_{k=1}^n c(x - X_{ks}), \quad (3.6)$$

gdzie  $c(u) = \frac{1}{2}[c^+(u) + c^-(u)]$  oznacza funkcję zliczającą:

$$c^+(u) = \begin{cases} 1, & \text{dla } u \geq 0, \\ 0, & \text{dla } u < 0, \end{cases}$$

$$c^-(u) = \begin{cases} 1, & \text{dla } u > 0, \\ 0, & \text{dla } u \leq 0. \end{cases}$$

Średnią dystrybuantę  $H(x)$  możemy teraz estymować jako łączny empiryczny rozkład

po wszystkich zmiennych losowych  $X_{11}, \dots, X_{nt}$ :

$$\hat{H}(x) = \frac{1}{t} \sum_{s=1}^t \hat{F}_s(x) = \frac{1}{t} \sum_{s=1}^t \frac{1}{n} \sum_{k=1}^n c(x - X_{ks}) = \frac{1}{N} \sum_{s=1}^t \sum_{k=1}^n c(x - X_{ks}). \quad (3.7)$$

Używając wzorów (3.6) i (3.7) można pokazać, że asymptotycznie nieobciążony estymator  $p_s$  jest równy:

$$\hat{p}_s = \int_{\mathbb{R}} \hat{H}(x) d\hat{F}_s(x) = \frac{1}{n} \sum_{k=1}^n \frac{1}{N} (R_{ks} - 0.5) = \frac{1}{N} (\bar{R}_{.s} - 0.5) \quad (3.8)$$

dla  $s = 1, \dots, k$  i  $N = n \cdot t$ .

Wyznaczenie estymatora  $\hat{p}_s$  umożliwi nam estymację nieznanych parametrów.

### 3.3 Testowanie hipotez

W tym rozdziale testujemy hipotezę zerową, że wszystkie obserwacje pochodzą z tego samego rozkładu przeciwko globalnej alternatywie. Problem testowania zapisujemy:

$$H_0 : F_1 = \dots = F_t,$$

$$H_1 : \exists_{i,j \in \{1, \dots, t\}} i \neq j \wedge F_i \neq F_j.$$

Przyjmijmy następujące oznaczenia: niech  $F = (F_1, \dots, F_t)^T$  będzie wektorem dystrybuant,  $p = (p_1, \dots, p_t)^T$  wektorem względnych wyników leczenia, a  $\hat{p} = (\hat{p}_1, \dots, \hat{p}_t)^T$  wektorem odpowiadających im estymatorów. Dodatkowo, niech  $\hat{V}_n$  oznacza macierz kowariancji rang a  $C$  odpowiednią macierz kontrastu.

#### Definicja 3.1

Macierz  $C$  nazywamy **macierzą kontrastu** jeżeli  $C_{r \times d} \mathbf{1}_d = \mathbf{0}_{r \times 1}$ , gdzie  $\mathbf{1}_d = (1, 1, \dots, 1)^T$  oznacza  $d$ -wymiarowy wektor złożony z 1.

Testowanie  $H_0$  przeciwko  $H_1$  można zapisać przy pomocy macierzy kontrastu jako:

$$H_0 : CF = 0,$$

$$H_1 : CF \neq 0.$$

Podamy rozwiązanie tego zagadnienia przy użyciu statystyki typu Walda i statystyki typu ANOVA.

#### 3.3.1 Statystyka typu Walda

Statystyka typu Walda (WTS) dana jest wzorem:

$$Q_n(C) = n\hat{p}^T C^T [C\hat{V}_n C^T]^+ C\hat{p}. \quad (3.9)$$

Notacja  $[C\hat{V}_n C^T]^+$  oznacza tutaj odwrotność typu Moore'a-Penrose'a macierzy  $C\hat{V}_n C^T$ .

Można pokazać (Brunner, 1997), że przy prawdziwości  $H_0$  statystyka  $Q_n(C)$  ma asymptotycznie rozkład chi-kwadrat z  $f = \text{rank}(C)$  stopniami swobody.

### 3.3.2 Statystyka typu ANOVA

Statystykę typu ANOVA (ATS) przedstawiamy jako:

$$A_n(C) = \frac{n}{\text{tr}(T\hat{V})} \hat{p}^T T \hat{p}, \quad (3.10)$$

gdzie  $T = C^T [CC^T]^- C$  a  $[CC^T]^-$  jest uogólnioną macierzą odwrotną  $CC^T$ .

Przy prawdziwości hipotezy zerowej statystykę  $A_n(C)$  można aproksymować rozkładem  $F_{(\hat{f}, \infty)}$ , gdzie  $\hat{f} = \frac{[\text{tr}(T\hat{V})]^2}{\text{tr}(T\hat{V}T\hat{V})}$  (Brunner, 1999).

Rozważania w tym rozdziale dotyczą przypomnianego wyżej modelu LD-F1. Dane dzielone są w nim na  $k = 1, \dots, n$  obserwacji i  $s = 1, \dots, t$  punktów czasu. W modelu zawierającym więcej czynników jak np. F1-LD-F1 niezbędne jest dodanie odpowiednich indeksów. Przykładowo, jeżeli jednostki podlegające eksperymentowi podzielone są na  $i = 1, \dots, a$  grup z  $j = 1, \dots, n_i$  jednostkami w  $i$ -tej grupie, to  $X_{ijs}$  oznacza zmienną losową dla  $j$ -tej jednostki w  $i$ -tej grupie w  $s$ -tym punkcie czasu.

## Rozdział 4

# Przykładowy eksperyment

W poprzednim rozdziale podaliśmy statystyki testowe dla danych longitudinalnych pochodzących z prostego modelu LD-F1. W tym modelu dane oznaczane są jako:  $X_{ks}$ , gdzie  $k$  oznacza numer obserwacji, a  $s$  jest punktem czasu. Przeprowadzane eksperymenty często wymagają budowania bardziej zaawansowanych modeli. Wprowadzane są wtedy różnorodne kryteria, które dzielą obserwacje na dodatkowe grupy. Przykładowo dla modelu dzielącego dane według czynnika A, przy danych pojawia się kolejny indeks  $i$ , oznaczający numer grupy, do której należą. Dane oznaczane są wtedy jako  $X_{iks}$ . Jednym z modeli wykorzystującym podział na grupy jest model F1-LD-F1. Użyjemy go do pokazania, jak wygląda testowanie przy bardziej skomplikowanym schemacie.

### 4.1 Model

Rozważmy model F1-LD-F1, gdzie czynnik A dzieli dane na dwie grupy, czyli  $i = 1, 2$ . Grupa pierwsza zawiera  $n_1$  a druga  $n_2$  obserwacji. Dane badane są w dwóch punktach czasu  $s$ , gdzie  $s = 1, 2$ . Łącznie model zawiera  $N = 2(n_1 + n_2)$  obserwacji.

		Dane		Rozkłady brzegowe	
		Czas		Czas	
Czynnik A	Obserwacja	$s = 1$	$s = 2$	$s = 1$	$s = 2$
$i = 1$	$k = 1$	$X_{111}$	$X_{112}$	$F_{11}$	$F_{12}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$k = n_1$	$X_{1n_11}$	$X_{1n_12}$	$F_{11}$	$F_{12}$
$i = 2$	$k = 1$	$X_{211}$	$X_{212}$	$F_{21}$	$F_{22}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$k = n_2$	$X_{2n_21}$	$X_{2n_22}$	$F_{21}$	$F_{22}$

Tabela 4.1: Model F1-LD-F1

Dane dla tego modelu to niezależne wektory losowe:

$$X_{ik} = (X_{ik1}, X_{ik2})' \text{ dla } k = 1, \dots, n_i.$$

O dystrybuantach brzegowych:

$$X_{iks} \sim F_{is}, \quad i = 1, 2, \quad s = 1, 2.$$



Testujemy, czy czas ma wpływ na wyniki leczenia, czyli, czy zachodzi efekt czasu (T). Interesuje nas także obecność efektu grupy (A) - czy między danymi z grupy numer 1 i danymi z grupy numer 2 widoczne są różnice? Sprawdźmy również jak wygląda interakcja między czasem a grupami (AT).

W rozdziale pierwszym wprowadziliśmy czynniki whole-plot i sub-plot. W powyższej tabeli czynnik whole-plot „A” obrazuje efekt grupy a czynnik sub-plot „Czas” reprezentuje efekt czasu. Poszczególne wiersze i kolumny tabeli reprezentują interakcje (AT).

## 4.2 Testowanie

Zapiszmy dane z modelu F1-LD-F1 z użyciem rang. W tym celu dokonujemy rangowania, czyli zastąpienia wszystkich obserwacji  $X_{iks}$  przez ich rangi  $R_{iks}$ . Zmienna  $\bar{R}_{i.s} = \frac{1}{n_i} \sum_{k=1}^{n_i} R_{iks}$  oznacza średnią rang w grupie  $i$  w punkcie czasu  $s$ .

		Dane		Średnia rang	
		Czas		Czas	
Czynnik A	Obserwacja	$s = 1$	$s = 2$	$s = 1$	$s = 2$
$i = 1$	$k = 1$	$R_{111}$	$R_{112}$	$\bar{R}_{1.1}$	$\bar{R}_{1.2}$
	$\vdots$	$\vdots$	$\vdots$		
	$k = n_1$	$R_{1n_11}$	$R_{1n_12}$		
$i = 2$	$k = 1$	$R_{211}$	$R_{212}$	$\bar{R}_{2.1}$	$\bar{R}_{2.2}$
	$\vdots$	$\vdots$	$\vdots$		
	$k = n_2$	$R_{2n_21}$	$R_{2n_22}$		

Tabela 4.2: Rangowanie dla danych z modelu F1-LD-F1

### 4.2.1 Efekt grupy

Badamy, czy podział danych na grupy według czynnika A ma wpływ na wyniki eksperymentu. Hipotezę zerową można zapisać za pomocą odpowiedniej macierzy kontrastu  $C_A = (1, 1, -1, -1)$  i wektora dystrybuant  $F$ :

$$H_0^A : C_A F = 0. \quad (4.1)$$

Po podstawieniu wartości otrzymujemy:

$$C_A F = (1, 1, -1, -1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \end{pmatrix} = F_{11} + F_{12} - F_{21} - F_{22} = \frac{1}{2}(\bar{F}_1 - \bar{F}_2) = 0. \quad (4.2)$$

A zatem równoważną formą zapisu hipotezy zerowej jest:

$$H_0^A : \bar{F}_1 = \bar{F}_2. \quad (4.3)$$

W rozdziale trzecim wyznaczyliśmy statystykę  $p_s$  i jej estymator  $\hat{p}_s = \frac{1}{N}(\bar{R}_{i.s} - 0.5)$  dla  $s = 1, 2$ ,  $i = 1, 2$ ,  $N = 2(n_1 + n_2)$ . Przy użyciu wektora estymatorów  $\hat{p}_s$ , macierzy  $\hat{V}_n$  i macierzy kontrastu  $C_A = (1, 1, -1, -1)$  możemy obliczyć wartość statystyki typu Walda i statystyki typu ANOVA oraz zweryfikować powyższą hipotezę.

### 4.2.2 Efekt czasu

W podobny sposób sprawdzimy, czy w analizowanym eksperymencie zachodzi efekt czasu. Do zapisania hipotezy zerowej korzystamy z macierzy kontrastu  $C_T = (1, -1, 1, -1)$  i wektora dystrybuant  $F$ :

$$H_0^T : C_T F = 0. \quad (4.4)$$

W rezultacie:

$$C_T F = (1, -1, 1, -1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \end{pmatrix} = F_{11} - F_{12} + F_{21} - F_{22} = \frac{1}{2}(\bar{F}_{.1} - \bar{F}_{.2}) = 0. \quad (4.5)$$

A zatem równoważną formą zapisu hipotezy zerowej jest:

$$H_0^T : \bar{F}_{.1} = \bar{F}_{.2}. \quad (4.6)$$

Przy użyciu wektora estymatorów  $\hat{p}$ , macierzy  $\hat{V}_n$  i macierzy kontrastu  $C_T = (1, -1, 1, -1)$  obliczamy wartość statystyki typu Walda i statystyki typu ANOVA i weryfikujemy  $H_0^T$ .

### 4.2.3 Interakcja

Rozważmy macierz kontrastu  $C_{AT} = (1, -1, -1, 1)$  i wektor dystrybuant  $F$ . Hipotezę zerową o braku interakcji zapisujemy jako:

$$H_0^{AT} : C_{AT} F = 0. \quad (4.7)$$

Liczmy:

$$C_{AT} F = (1, -1, -1, 1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \end{pmatrix} = F_{11} - F_{12} - F_{21} + F_{22} = 0. \quad (4.8)$$

Hipotezę zerową można zatem sformułować przez:

$$H_0^{AT} : F_{11} - F_{12} = F_{21} - F_{22}. \quad (4.9)$$

Testujemy z użyciem statystyk WTS i ATS wyliczonych dla wektora estymatorów  $\hat{p}$ , macierzy  $\hat{V}_n$  i macierzy kontrastu  $C_{AT} = (1, -1, -1, 1)$ .

W kolejnym rozdziale zajmiemy się praktyczną analizą danych. Zbudujemy odpowiednie modele i wykresy dla danych, wyliczymy wartości statystyk WTS i ATS oraz przetestujemy odpowiednie hipotezy. Analizy wykonamy za pomocą pakietu `nparLD` w środowisku `R`.

## Rozdział 5

# Przykłady

W ostatnim rozdziale zajmiemy się analizą danych longitudinalnych w praktyce. Za pomocą przykładów pokażemy, w jaki sposób można badać różnorodne eksperymenty czynnikowe. Skorzystamy ze zbudowanego w tym celu pakietu `nparLD` i zawartych w nim funkcji. Będą to między innymi:

- `ld.f1()`, `ld.f2()`, `f1.ld.f1()`, `f2.ld.f1()`, `f1.ld.f2()` - funkcje tworzące wybrany model dla danych,
- `nparLD()` - funkcja automatycznie identyfikująca najlepszy model dla danych,
- `print()` - funkcja wyświetlająca wyniki oraz wartości statystyk WTS i ATS,
- `summary()` - funkcja wyświetlająca dodatkowe informacje o modelu,
- `plot()` - funkcja tworząca wykresy dla danych i dla modelu.

### 5.1 Badanie ze stomatologii

Autorzy badania:	R. Potthoff, S. Roy (1964)
Opis badania:	Badamy odległość między środkiem przysadki a szczeliną skrzydłowo-szczękową u 16 chłopców. Dane gromadzimy czterokrotnie - gdy chłopcy mają: 8, 10, 12 i 14 lat.
Model dla danych:	LD-F1
Hipoteza:	Testujemy hipotezę o braku efektu czasu: $H_0^T : F_8 = F_{10} = F_{12} = F_{14}$ , gdzie $F_s$ jest dystrybuantą brzegową badanej odległości, gdy chłopcy są w wieku $s$ lat.

#### 5.1.1 Analiza nieparametryczna

Przeanalizujemy kod w R napisany dla tego zadania.

Wczytanie biblioteki `nparLD` oraz danych ze zbioru `dental`:

```
> library("nparLD")
> data("dental")
```

Budowa modelu:

```
> model <- nparLD(resp ~ time, data = dental, subject = "subject")
```

```
Total number of observations: 64
Total number of subjects: 16
Total number of missing observations: 0
```

```
LD F1 Model
```

```
-----
Check that the order of the time level is correct.
```

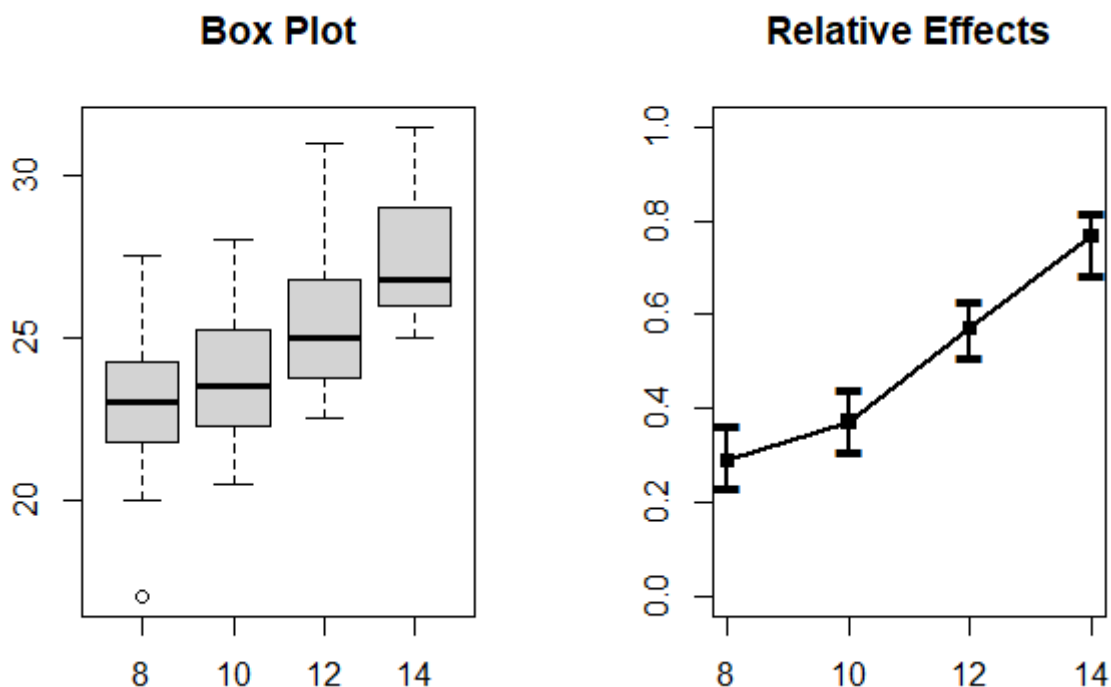
```
Time level: 8 10 12 14
```

```
If the order is not correct, specify the correct order in time.order.
```

Zauważmy, że funkcja `nparLD()` sama dobrała najlepszy model do analizowanych danych i automatycznie wyświetliła informacje o jego strukturze.

Wyświetlenie wykresów:

```
> par(mfrow=c(1,2))
> boxplot(resp ~ time, data = dental,
+         xlab = "time",ylab="", main = "Box Plot")
> plot(model)
```



Rysunek 5.1: Wykres pudełkowy i 95% przedziały ufności dla  $p_s$  w badaniu ze stomatologii

Na ilustracji po lewej stronie widzimy cztery wykresy pudełkowe. Dla każdego z pomiarów w punkcie czasu  $s = 8, 10, 12, 14$  zaznaczone są: minimum, pierwszy kwartył, mediana, trzeci kwartył, maksimum oraz rozstęp międzykwartyłowy IQR. Przeanalizujmy ogólną strukturę każdego z boxplotów. Wykres dla  $s = 8$  jest prawie symetryczny, co świadczy o równomiernym rozłożeniu danych. Wraz ze wzrostem  $s$  zwiększa się skośność wykresów, a więc rozkład obserwacji jest coraz bardziej nieregularny. Możemy zatem wnioskować o istnieniu efektu

czasu. Rosnące wartości kwantyli także wskazują na wpływ czasu na pomiary.

Wykres po prawej stronie obrazuje 95% przedziały ufności dla statystyki  $p_s$ . Dla każdego z punktów pomiaru przedstawione są: dolna granica przedziału, obliczona wartość estymatora statystyki  $p_s$  i górna granica przedziału. Wartość estymatora rośnie wraz z upływem czasu, co oznacza, że badana odległość zwiększa się wraz z wiekiem chłopców.

Wyświetlenie szczegółowych informacji o przedziałach ufności:

```
> plot(model)$Conf.Int
```

	Group	Time	Nobs	RankMeans	RTE	Bias	Variance	Lower	Upper
1	Group1	8	16	19.0625	0.2900	-0.0048	0.0177	0.2340	0.3642
2	Group1	10	16	24.3125	0.3721	-0.0024	0.0190	0.3097	0.4436
3	Group1	12	16	37.0312	0.5708	0.0021	0.0154	0.5087	0.6291
4	Group1	14	16	49.5938	0.7671	0.0051	0.0177	0.6846	0.8175

Dla każdego z punktów  $s = 8, 10, 12, 14$  tabela wyświetla: średnią rang (RankMeans), wskaźnik względnego efektu czasu (RTE), obciążenie estymatora (Bias), wariancję estymatora (Variance) oraz wartość dolnej (Lower) i górnej (Upper) granicy przedziału ufności.

Wielkość RTE można interpretować jako prawdopodobieństwo, że losowo wybrana obserwacja z całego zbioru danych ma mniejszą wartość, niż losowo wybrana obserwacja z analizowanej grupy. Przykładowo  $RTE = 0.3721$  dla  $s = 10$  oznacza, że losowo wybrana zmienna ze zbioru `dental` z prawdopodobieństwem równym 37.21% ma wartość mniejszą, niż losowa obserwacja z grupy, gdzie  $s = 10$ . Wiemy, że  $\hat{p}_8 < \hat{p}_{10} < \hat{p}_{12} < \hat{p}_{14}$ . Widzimy, że między wartościami RTE zachodzi analogiczna nierówność:  $RTE_8 < RTE_{10} < RTE_{12} < RTE_{14}$ .

Podsumowanie informacji o danych:

```
> summary(model)
```

Model:

LD F1 Model

Call:

```
resp ~ time
```

Relative Treatment Effect (RTE):

	RankMeans	Nobs	RTE
time8	19.06250	16	0.2900391
time10	24.31250	16	0.3720703
time12	37.03125	16	0.5708008
time14	49.59375	16	0.7670898

Wald-Type Statistic (WTS):

	Statistic	df	p-value
time	94.47718	3	2.391503e-20

ANOVA-Type Statistic (ATS):

```

Statistic      df      p-value
time  31.48774  2.700785  1.437729e-18

```

Dla każdej grupy pomiarów w punkcie  $s = 8, 10, 12, 14$  wyświetlone są: liczba obserwacji (Nobs), średnia rang (RankMeans) oraz wskaźnik względnego efektu czasu (RTE).

Funkcja `summary()` podaje także obliczone wartości statystyk typu Walda (Wald-Type Statistic WTS) i typu ANOVA (ANOVA-Type Statistic ATS), których użyjemy do przetestowania hipotezy  $H_0^T$  o braku efektu czasu. W obu przypadkach  $p$ -wartość (p-value) jest mniejsza od przyjętego poziomu istotności  $\alpha = 0.001$ . Mamy zatem 99.999% pewności, że hipoteza zerowa jest fałszywa, co implikuje, że dla analizowanych danych zachodzi efekt czasu.

Przy pomocy poprawki Bonferroniego badamy różnicę między dystrybuantami  $F_s$ :

```

> m8vs10 <- which(((dental$time == 8) + (dental$time == 10)) == 1)
> m8vs12 <- which(((dental$time == 8) + (dental$time == 12)) == 1)
> m8vs14 <- which(((dental$time == 8) + (dental$time == 14)) == 1)
> model8vs10 <- nparLD(resp ~ time, data = dental[m8vs10,],
+                      subject = "subject", description = FALSE)
> model8vs12 <- nparLD(resp ~ time, data = dental[m8vs12,],
+                      subject = "subject", description = FALSE)
> model8vs14 <- nparLD(resp ~ time, data = dental[m8vs14,],
+                      subject = "subject", description = FALSE)

```

W tabeli przedstawimy  $p$ -wartości dla ATS zwrócone przez `summary` oraz  $p$ -wartości powstałe po korekcie Bonferroniego ( $p$ -wartość po korekcie Bonferroniego jest obliczana przez trzykrotne zwiększenie oryginalnej  $p$ -wartości):

Porównanie	Hipoteza	$p$ -wartość	$p$ -wartość po korekcie
$s = 8$ przeciwko $s = 10$	$H_0^T : F_8 = F_{10}$	0.2204	0.6612
$s = 8$ przeciwko $s = 12$	$H_0^T : F_8 = F_{12}$	$1.7162 \times 10^{-09}$	$5.1487 \times 10^{-09}$
$s = 8$ przeciwko $s = 14$	$H_0^T : F_8 = F_{14}$	$3.1572 \times 10^{-16}$	$9.4716 \times 10^{-16}$

Tabela 5.1: Porównanie dystrybuant dla różnych grup pomiarowych w badaniu ze stomatologii.

Dla  $p$ -wartości równej 0.6612 nie mamy podstaw do odrzucenia hipotezy zerowej na poziomie istotności  $\alpha = 0.001$ . Wnioskujemy, że dystrybuanta  $F_8$  jest równa dystrybuancie  $F_{10}$ . W pozostałych przypadkach odrzucamy hipotezę zerową - dystrybuanty nie są sobie równe. Zauważmy, że  $p$ -wartości zmniejszają się wraz ze wzrostem  $s$ ,  $s = 10, 12, 14$ . Wskazuje to na fakt, że analizowany dystans między środkiem przysadki a szczeliną skrzydłowo-szczękową zwiększa się wraz z dorastaniem chłopców. Dodatkowo, różnice w  $p$ -wartościach wskazują na niezgodności między rozkładami dla grup wiekowych:  $s = 8$  i  $s = 12$  oraz  $s = 8$  i  $s = 14$ .

### 5.1.2 Analiza parametryczna

Przeanalizujmy dane ponownie, tym razem z użyciem metod parametrycznych.

Wczytanie biblioteki i konstrukcja modelu:

```
> library("nlme")
> model <- lme(resp ~ time, data = dental, random = ~ 1 | subject)
```

Podsumowanie informacji o danych:

```
> summary(model)
```

```
Fixed effects: resp ~ time
                Value Std.Error DF   t-value p-value
(Intercept) 16.340625  1.1287202 47  14.477126     0
time          0.784375  0.0938154 47   8.360838     0
```

W tym przypadku także możemy odrzucić hipotezę zerową o braku efektu czasu, ponieważ  $p$ -wartość  $\sim 0$  jest mniejsza od poziomu istotności  $\alpha = 0.001$ .

Przy pomocy poprawki Bonferroniego badamy różnicę między dystrybuantami  $F_s$ :

```
> model8vs10 <- lme(resp ~ time, data = dental[m8vs10,], random = ~ + 1 | subject)
> model8vs12 <- lme(resp ~ time, data = dental[m8vs12,], random = ~ + 1 | subject)
> model8vs14 <- lme(resp ~ time, data = dental[m8vs14,], random = ~ + 1 | subject)
```

Ponownie porównujemy  $p$ -wartości:

Porównanie	Hipoteza	$p$ -wartość	$p$ -wartość po korekcie
$s = 8$ przeciwko $s = 10$	$H_0^T : F_8 = F_{10}$	0.1465	0.4395
$s = 8$ przeciwko $s = 12$	$H_0^T : F_8 = F_{12}$	$3 \times 10^{-04}$	$9 \times 10^{-04}$
$s = 8$ przeciwko $s = 14$	$H_0^T : F_8 = F_{14}$	0	0

Tabela 5.2: Porównanie dystrybuant w modelu parametrycznym w badaniu ze stomatologii.

W tym przypadku także nie mamy podstaw do odrzucenia hipotezy zerowej  $H_0^T : F_8 = F_{10}$ . W pozostałych przypadkach ponownie przyjmujemy hipotezę alternatywną, że badane dystrybuanty są różne.

Podsumowując, wnioski z analizy danych metodami parametrycznymi i nieparametrycznymi są podobne. Spowodowane jest to względnie symetrycznym rozkładem danych. Wprawdzie na wykresach pudełkowych zauważalna była asymetria, ale stopień skośności był lekki.

## 5.2 Badanie rozwoju szczura

Autorzy badania:	G.E.P. Box (1950), R.D. Wolfinger (1996)
Opis badania:	Przedmiotem badania jest waga 27 szczurów. Szczury zostały losowo podzielone na grupy o liczebnościach: 10, 7, 10. Pierwsza grupa była grupą kontrolną. W pozostałych dwóch do wody pitnej szczurów dodawane były odpowiednio: tyroksyna i tiouracyl. Szczury ważone były przed rozpoczęciem badania oraz regularnie, przez cztery tygodnie jego trwania.
Model dla danych:	F1-LD-F1
Hipoteza:	Testujemy hipotezę o <i>efekcie wachlarzowania</i> (fanning effect) zaproponowaną przez R.D. Wolfingera. Dokładniej, badamy wzrost zmienności danych w czasie.

Analizę przeprowadzimy tylko przy pomocy metod nieparametrycznych. Zaczniemy od konstrukcji modelu i wyświetlenia wykresów pudełkowych:

```
> plot(model)
> model <- nparLD(resp ~ time * group, data = rat,
+               subject = "subject", description = FALSE)
> par(mfrow=c(1,2))
> boxplot(resp ~ group * time, data = rat, names = FALSE,
+         col = c("grey", 2, 3), lwd = 2, main = "Box Plots", ylab="")
> axis(1, at = 2, labels = "0", cex = 2)
> axis(1, at = 5, labels = "1", cex = 2)
> axis(1, at = 8, labels = "2", cex = 2)
> axis(1, at = 11, labels = "3", cex = 2)
> axis(1, at = 14, labels = "4", cex = 2)
> legend(2, 196, c("Control", "Thiour", "Thyrox"), lwd = c(3, 3, 3),
+       col = c("grey", 2, 3), cex = 1)
> plot(model)
```

F1 LD F1 Model

-----

Check that the order of the time and group levels are correct.

Time level: 0 1 2 3 4

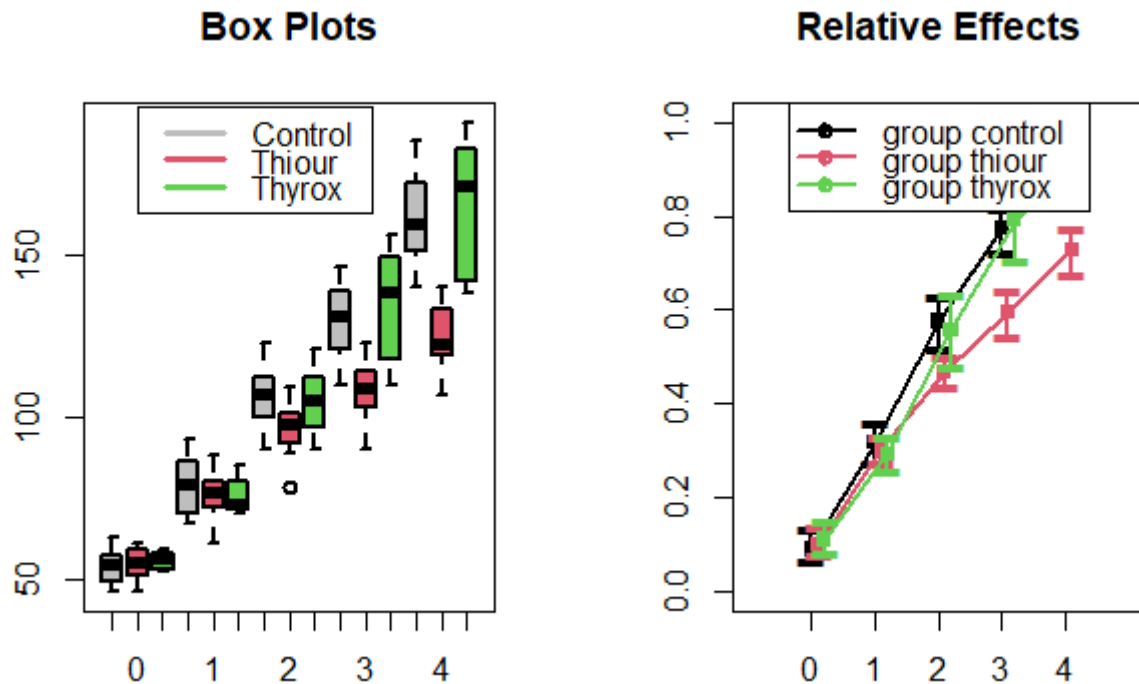
Group level: control thyrox thiour

If the order is not correct, specify the correct order  
in time.order or group.order.

Rysunek (5.2) przedstawia wykresy pudełkowe i 95% przedziały ufności dla względnych efektów leczenia. Na wykresach uwzględnione są: podział zwierząt na trzy grupy i pięć punktów czasu, w których dokonywano ważenia. Wykresy pudełkowe mają duży współczynnik skośności, co świadczy o nierównomiernym rozkładzie danych. Z wykresu dla 95% przedziałów ufności widać, że waga szczurów w każdej z grup zwiększa się wraz z upływem czasu.

Głównym celem doświadczenia jest odpowiedź na pytanie, czy istnieje interakcja między leczeniem a upływem czasu. Jeżeli czas i leczenie byłyby od siebie niezależne, to profile czasowe dla trzech grup byłyby równoległe. Wykresy z rysunku (5.2) przeczą przypuszczeniu





Rysunek 5.2: Wykres pudełkowy i 95% przedział ufności dla  $p_s$  w badaniu rozwoju szczura.

o niezależności tych czynników. Widzimy, że linia dla grupy z tiouracylem nie jest równoległa ani do linii dla tyroksyny, ani dla grupy kontrolnej.

Przetestujemy hipotezę o braku interakcji między czasem a efektami leczenia.

```
> summary(model)
```

Model:

F1 LD F1 Model

Call:

```
resp ~ time * group
```

Relative Treatment Effect (RTE):

	RankMeans	Nobs	RTE
groupcontrol	72.92000	50	0.53644444
groupthyrox	72.67143	35	0.53460317
groupthiour	59.81000	50	0.43933333
time0	14.20714	27	0.10153439
time1	41.55714	27	0.30412698
time2	72.57857	27	0.53391534
time3	97.68810	27	0.71991182
time4	116.30476	27	0.85781305
groupcontrol:time0	12.75000	10	0.09074074
groupcontrol:time1	43.30000	10	0.31703704
groupcontrol:time2	78.25000	10	0.57592593
groupcontrol:time3	105.00000	10	0.77407407

groupcontrol:time4	125.30000	10	0.92444444
groupthyrox:time0	15.57143	7	0.11164021
groupthyrox:time1	40.07143	7	0.29312169
groupthyrox:time2	75.78571	7	0.55767196
groupthyrox:time3	107.21429	7	0.79047619
groupthyrox:time4	124.71429	7	0.92010582
groupthiour:time0	14.30000	10	0.10222222
groupthiour:time1	41.30000	10	0.30222222
groupthiour:time2	63.70000	10	0.46814815
groupthiour:time3	80.85000	10	0.59518519
groupthiour:time4	98.90000	10	0.72888889

Wald-Type Statistic (WTS):

	Statistic	df	p-value
group	12.52657	2	1.904977e-03
time	3619.03739	4	0.000000e+00
group:time	70.34311	8	4.199050e-12

ANOVA-Type Statistic (ATS):

	Statistic	df	p-value
group	5.286582	1.922792	5.654723e-03
time	1008.512138	1.990411	0.000000e+00
group:time	11.093940	3.516933	3.616929e-08

Modified ANOVA-Type Statistic for the Whole-Plot Factors:

	Statistic	df1	df2	p-value
group	5.286582	1.922792	19.23468	0.01563658

Testujemy hipotezę o braku interakcji na poziomie istotności  $\alpha = 0.001$ .  $p$ -wartość testu Walda wynosi  $4.199 \times 10^{-12}$ . Test ANOVA daje  $p$ -wartość równą  $3.6169 \times 10^{-8}$ . Obie wielkości są znacznie mniejsze od  $\alpha$  więc możemy odrzucić hipotezę zerową z prawie 100% pewnością. Możemy także badać istnienie interakcji w konkretnych punktach czasu. W tym celu dokonujemy porównań między wagą szczurów przed rozpoczęciem badania, a wagą odpowiednio po 1, 2 i kolejnym tygodniu.

W celu kontroli błędu pierwszego rodzaju możemy wprowadzić korektę Bonferroniego - analogicznie jak we wcześniej analizowanym badaniu ze stomatologii.

## 5.3 Badanie zaburzeń oddychania

Autorzy badania:	G.G. Koch, G.J. Carr, I.A. Amara, M.E. Stokes, T.J. Uryniak (1990)
Opis badania:	Analizujemy badanie kliniczne dla pacjentów z zaburzeniami oddychania. Pacjenci pochodzili z dwóch niezależnych centrów badawczych, ich łączna liczba wynosiła 111. Każdego z uczestników badania przydzielono losowo do grupy z aktywnym leczeniem (grupa aktywna) albo do grupy z leczeniem placebo (grupa placebo). Status każdego z pacjentów oceniono według skali: 0 = okropny, 1 = słaby, 2 = znośny, 3 = dobry, 4 = wspaniały. Oceny dokonano przed rozpoczęciem badania oraz kolejno podczas czterech wizyt kontrolnych.
Model dla danych:	F2-LD-F1
Hipoteza:	Badamy, czy wyniki dla grupy aktywnej (A) są lepsze niż dla grupy placebo (P). Analizujemy też ogólne wyniki leczenia z uwzględnieniem centrum badawczego, z którego pochodzą pacjenci.

Dane analizujemy w sposób nieparametryczny z użyciem rangowania. Jedną z zalet takiego podejścia jest łatwa obsługa danych porządkowych niewymagająca dodatkowych transformacji. Standardowo zaczniemy od konstrukcji modelu i analizy wykresów:

```
> library("nparLD")
> data("respiration")
> model <- nparLD(resp ~ time * center * treatment, data = respiration,
+               subject = "patient", description = FALSE)

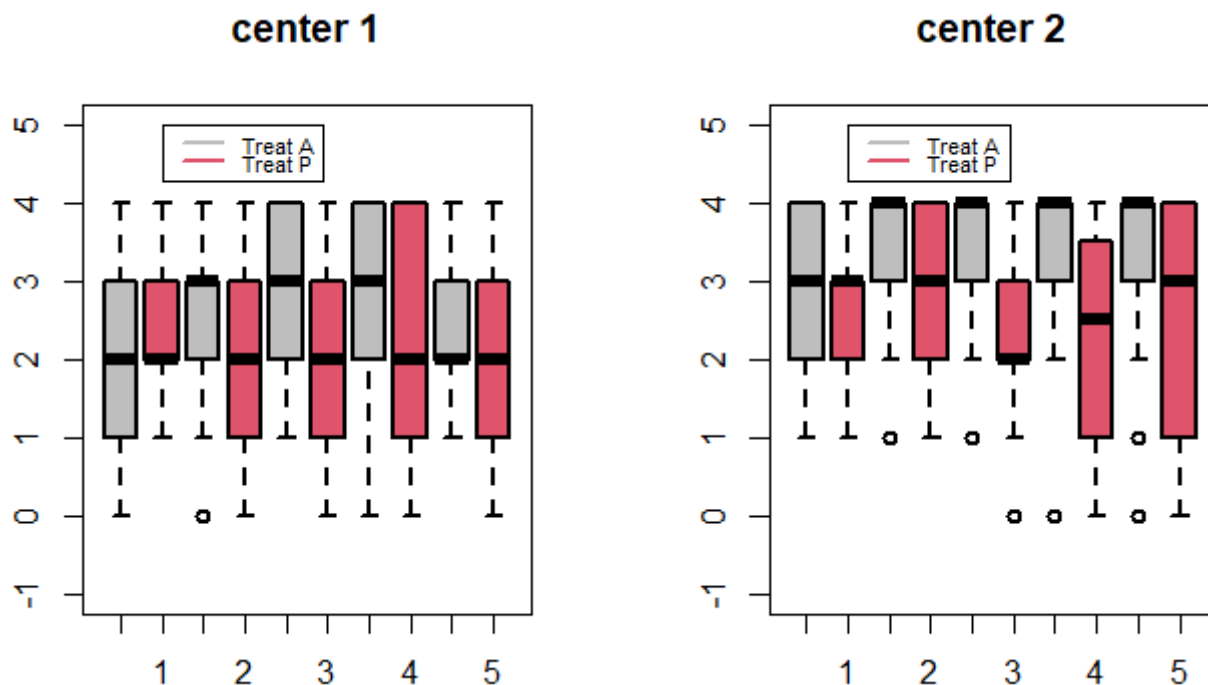
F2 LD F1 Model
-----
Check that the order of the time, group1, and group2 levels are correct.
Time level:   1 2 3 4 5
Group1 level:  1 2
Group2 level:  A P
If the order is not correct, specify the correct order in time.order, group1.order,
or group2.order.

> par(mfrow = c(1, 2))
> center <- respiration[, "center"]
labs = c("", "1", "", "2", "", "3", "", "4", "", "5")
boxplot(resp ~ treatment * time, data = respiration[which(center == 1), ],
        names = labs, ylim = c(-1, 5), col = c("grey", 2), lwd = 2,
        main = "center 1", ylab = "", xlab = "")
legend(2, 5, c("Treat A", "Treat P"), lwd = c(2, 2), col = c("grey", 2),
      cex = 0.7)
boxplot(resp ~ treatment * time, data = respiration[which(center == 2), ],
        names = labs, ylim = c(-1, 5), col = c("grey", 2), lwd = 2,
```

```

main = "center 2", xlab = "", ylab = "")
legend(2, 5, c("Treat A", "Treat P"), lwd = c(2, 2), col = c("grey", 2),
      cex = 0.7)
> plot(model)

```



Rysunek 5.3: Wykresy pudełkowe dla badania zaburzeń oddychania

Przeanalizujemy wykresy pudełkowe widoczne na rysunku (5.3). Widzimy, że wartość dolnego kwartyla w grupie aktywnej jest większa niż w grupie placebo; ogólnie wyniki dla grupy A są lepsze niż dla grupy P. Zauważalne są także różnice w rezultatach dla centrum 1 i dla centrum 2.

Rysunek (5.4) przedstawia 95% przedziały ufności dla statystyk  $p_{ijs}$ . W obu centach badawczych wyniki dla grupy placebo pozostają na stałym poziomie bliskim 0.4. W centrum 1 widoczna jest zmienność w leczeniu dla grupy A. Wartość  $p_{11s}$  rośnie dla czasów 1 – 3, następnie utrzymuje się na stałym poziomie i maleje przy ostatnim badaniu w czasie 5. W drugim centrum wartość  $p_{21s}$  rośnie między 1 – 2, po czym pozostaje stała.

Wyświetlmy teraz podsumowanie informacji o danych i przeanalizujemy występowanie efektu czasu, efektu podziału na centra badawcze oraz istnienie interakcji między nimi.

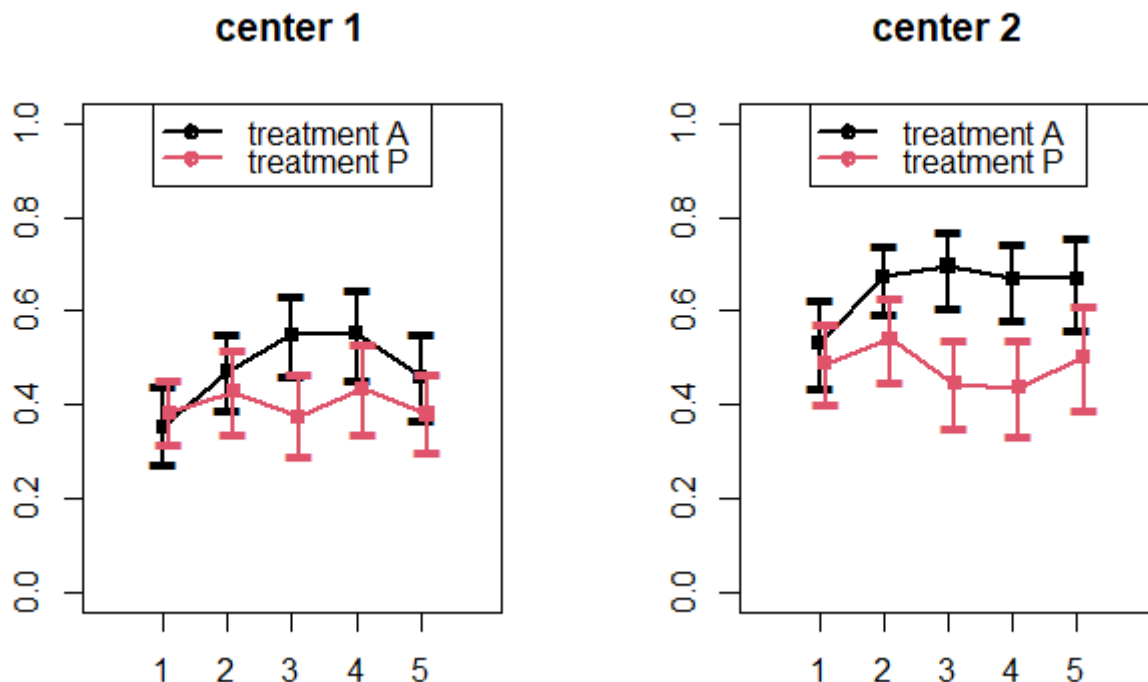
```
> print(model)
```

```
Model:
F2 LD F1 Model
```

```
Call:
resp ~ time * center * treatment
```

```
Wald-Type Statistic (WTS):
```

Statistic	df	p-value
-----------	----	---------

Rysunek 5.4: 95% przedziały ufności statystyki  $p_{ijs}$  dla badania zaburzeń oddychania

```

center          10.2569587  1 0.001361700
treatment       9.3451482  1 0.002235766
time           17.4568433  4 0.001575205
center:treatment  1.2365618  1 0.266134717
center:time      8.7200395  4 0.068491057
treatment:time  17.5434583  4 0.001515158
center:treatment:time 0.2898785  4 0.990458142

```

ANOVA-Type Statistic (ATS):

	Statistic	df	p-value
center	10.25695866	1.000000	0.0013616998
treatment	9.34514819	1.000000	0.0022357657
time	4.43527016	3.320559	0.0028528788
center:treatment	1.23656176	1.000000	0.2661347165
center:time	1.60699585	3.320559	0.1802120504
treatment:time	5.46185031	3.320559	0.0005867392
center:treatment:time	0.05915234	3.320559	0.9866660535

Modified ANOVA-Type Statistic for the Whole-Plot Factors:

	Statistic	df1	df2	p-value
center	10.256959	1	104.9255	0.001803091
treatment	9.345148	1	104.9255	0.002836284
center:treatment	1.236562	1	104.9255	0.268676117

Funkcja `print()` automatycznie wyświetliła wyniki dla wszystkich zmiennych w modelu oraz interakcji między nimi. Wyliczone zostały też wartości statystyk WTS, ATS i zmodyfikowana

statystyka ATS dla czynników whole-plot (`center`, `treatment` i interakcja między nimi).

Testujemy hipotezę zerową o braku interakcji między leczeniem a upływem czasu.  $p$ -wartości obliczone dla WTS i ATS wynoszą odpowiednio: 0.0015 i 0.0006. Obie  $p$ -wartości są mniejsze od przyjętego poziomu istotności  $\alpha = 0.05$ , co świadczy o istnieniu interakcji między zmiennymi. Do podobnych wniosków doszliśmy analizując wykresy dla 95% przedziałów ufności na rysunku (5.4).

Ze zmodyfikowanej ATS dostajemy informacje o istotności podziału na centrum ( $p$ -wartość  $0.0018 < 0.05$ ) oraz dla istnienia wyraźnych efektów leczenia ( $p$ -wartość  $0.0028 < 0.05$ ).

Przeanalizujemy teraz dane za pomocą metod parametrycznych i porównajmy wyniki.

```
> library("nlme")
> model <- lme(resp ~ time * treatment * center, data = respiration,
+             random = ~ 1 | patient)
> summary(model)
```

Linear mixed-effects model fit by REML

```
Data: respiration
      AIC      BIC    logLik
1585.142 1628.187 -782.5712
```

Random effects:

```
Formula: ~1 | patient
      (Intercept) Residual
StdDev:  0.8673626 0.8073981
```

Fixed effects: resp ~ time \* treatment \* center

	Value	Std.Error	DF	t-value	p-value
(Intercept)	1.3148148	0.5216425	440	2.5205287	0.0121
time	0.1740741	0.1098730	440	1.5843212	0.1138
treatmentP	0.3729931	0.7261302	107	0.5136724	0.6085
center	0.8111111	0.3299157	107	2.4585408	0.0156
time:treatmentP	-0.1399608	0.1529440	440	-0.9151113	0.3606
time:center	-0.0407407	0.0694898	440	-0.5862840	0.5580
treatmentP:center	-0.2782293	0.4604257	107	-0.6042872	0.5469
time:treatmentP:center	-0.0209588	0.0969789	440	-0.2161167	0.8290

Correlation:

	(Intr)	time	trtmnP	center	tm:trP	tm:cnt	trtmP:
time		-0.632					
treatmentP		-0.718	0.454				
center		-0.949	0.599	0.682			
time:treatmentP		0.454	-0.718	-0.632	-0.431		
time:center		0.599	-0.949	-0.431	-0.632	0.682	
treatmentP:center		0.680	-0.430	-0.948	-0.717	0.599	0.453
time:treatmentP:center		-0.430	0.680	0.599	0.453	-0.948	-0.717

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-3.02642169	-0.56806580	0.05714321	0.65231170	2.80043199

Number of Observations: 555

Number of Groups: 111

Zauważmy, że wnioski wyciągnięte przy pomocy metody parametrycznej (`lme()`) są całkowicie różne, od tych otrzymanych metodą nieparametryczną (`nparLD()`).

Testujemy hipotezę zerową o tym, że dana zmienna jest istotna. Rozważymy pięć problemów testowania dla efektu czasu, leczenia i centrum oraz dla interakcji czasu i leczenia oraz interakcji leczenia i centrum. Tabela (5.3) przedstawia porównanie  $p$ -wartości dla zmiennych z modelu skonstruowanego obiema metodami. Zwróćmy uwagę na to, że tylko w 40% otrzymaliśmy takie same rezultaty. Sprzeczność wyników ma związek ze strukturą danych. Informacje od pacjentów zbierane są w sposób uporządkowany z podziałem na kategorie - w takim przypadku metody parametryczne nie powinny być stosowane.

zmienna	nparLD()		lme()	
	$p$ -wartość ATS	istotność	$p$ -wartość	istotność
czas	0.0028	tak	0.1138	nie
leczenie	0.0022	tak	0.6085	nie
centrum	0.0013	tak	0.0156	tak
czas:leczenie	0.0006	tak	0.3606	nie
leczenie:centrum	0.2661	nie	0.5469	nie

Tabela 5.3: Testowanie hipotez metodą parametryczną i nieparametryczną

# Spis tabel

2.1	Model LD-F1 . . . . .	8
2.2	Model F1-LD-F1 . . . . .	9
2.3	Model F2-LD-F1 . . . . .	10
3.1	Model LD-F1 . . . . .	11
3.2	Rangowanie danych z modelu LD-F1 . . . . .	13
4.1	Model F1-LD-F1 . . . . .	16
4.2	Rangowanie dla danych z modelu F1-LD-F1 . . . . .	17
5.1	Porównanie dystrybuant dla różnych grup pomiarowych w badaniu ze stomatologii. . . . .	22
5.2	Porównanie dystrybuant w modelu parametrycznym w badaniu ze stomatologii. . . . .	23
5.3	Testowanie hipotez metodą parametryczną i nieparametryczną . . . . .	31



# Spis rysunków

3.1	Związek statystyki $p$ z dystrybucją $F$ . . . . .	13
5.1	Wykres pudełkowy i 95% przedziały ufności dla $p_s$ w badaniu ze stomatologii	20
5.2	Wykres pudełkowy i 95% przedział ufności dla $p_s$ w badaniu rozwoju szczura.	25
5.3	Wykresy pudełkowe dla badania zaburzeń oddychania . . . . .	28
5.4	95% przedziały ufności statystyki $p_{ijs}$ dla badania zaburzeń oddychania . . .	29

# Literatura

- [1] M.G. Akritas, S.F. Arnold, *Fully Nonparametric Hypotheses for Factorial Designs I: Multivariate Repeated Measures Designs*, Journal of the American Statistical Association (1994), 89, 336–343.
- [2] M.G. Akritas, S.F. Arnold, E. Brunner, *Nonparametric Hypotheses and Rank Statistics for Unbalanced Factorial Designs*, Journal of the American Statistical Association (1997), 92, 258–265.
- [3] M. G. Akritas, E. Brunner, *A Unified Approach to Rank Tests for Mixed Models*, Journal of Statistical Planning and Inference (1997), 61, 249–277.
- [4] B. Bandelow, E. Brunner, A. Broocks, D. Beinroth, G. Hajak, L. Pralle, E. Ruther, *The Use of the Panic and Agoraphobia Scale in a Clinical Trial*, Psychiatry Research (1998), 77, 43–49.
- [5] G.E.P Box, *Problems in the Analysis of Growth and Wear Curves*, Biometrics (1950), 6, 362–389.
- [6] E. Brunner, H. Dette, A. Munk, *Box-Type Approximations in Nonparametric Factorial Designs*, Journal of the American Statistical Association (1997), 92, 1494–1502.
- [7] E. Brunner, S. Domhof, F. Langer, *Nonparametric Analysis of Longitudinal Data in Factorial Experiments*, John Wiley & Sons, New York (2002).
- [8] E. Brunner, U. Munzel, M. L. Puri, *Rank-Score Tests in Factorial Designs with Repeated Measures*, Journal of Multivariate Analysis (1999), 70, 286–317.
- [9] E. Brunner, M. L. Puri, *Nonparametric Methods in Factorial Designs*, Statistical Papers (2001), 42, 1–52.
- [10] P.J. Diggle, K.Y. Liang, S.L. Zeger, *Analysis of Longitudinal Data*, Oxford University Press (1994).
- [11] G.G. Koch, G.J. Carr, I.A. Amara, M.E. Stokes, T.J. Uryniak, *Categorical Data Analysis* (1990).
- [12] F. Konietzschke, A.C. Bathke, L.A. Hothorn, E. Brunner, *Testing and Estimation of Purely Nonparametric Effects in Repeated Measures Designs*, Computational Statistics & Data Analysis (2010), 54, 1895–1905.
- [13] K. Noguchi, Y. R. Gel, E. Brunner, F. Konietzschke, *nparLD: An R Software Package for the Nonparametric Analysis of Longitudinal Data in Factorial Experiments*, Journal of Statistical Software (2012), 50, 12.

- [14] K. Noguchi, M. Latif, K. Thangavelu, Y. R. Gel, E. Brunner, *Nonparametric Analysis of Longitudinal Data in Factorial Experiments* (2022).
- [15] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Development Core Team (2012), *nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-104*, URL: <http://CRAN.R-project.org/package=nlme>.
- [16] R.F. Potthoff, S.N. Roy, *Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems*, *Biometrika* (1964), 51, 313–326.
- [17] R Development Core Team, *R: A Language and Environment for Statistical Computing*, Vienna, Austria (2012), ISBN 3-900051-07-0, URL: <http://www.R-project.org/>.
- [18] T. Rodig, M. Sedghi, F. Konietzschke, K. Lange, D. Ziebolz, M. Hulsmann, *Efficacy of Syringe Irrigation, RinsEndo® and Passive Ultrasonic Irrigation in Removing Debris from Irregularities in Root Canals with Different Apical Sizes*, *International Endodontic Journal* (2010), 43, 581–589.
- [19] SAS Institute Inc, *The SAS System, Version 9.1. SAS Institute Inc, Cary, NC* (2003), URL: <http://www.sas.com/>.