**СТАТИСТИКА БОЛЬШИХ ДАННЫХ – ИДЗ 2**

Студент: Эспинола Ривера, Хольгер Элиас

Тема: Статистический анализ динамики изменений наблюдаемой характеристики с течением времени

Для когорты индивидов проведены измеения исследуемой характеристики Y в моменты времени t1, …, td (Visit), расстояние между которыми одинакого. Помимо исследуемой характеристики измерялись сопутствующие факторы А (не меняющаяся со времени характеристика, измеряемая в момент начала исследования) и В (меняющаяся со временем характеристика, измерение которой проводилось в каждой временой точке). Из таблицы данных следует выбрать значения описанных характеристики, с персональным значением переменной “Variant”.

**Вопросы**

**1. Загрузить данные и отфильтрировать строки с персональным значением переменной “Variant”.**

Применение филтра для выбора данных для варианта 9:

> # collect the data which corresponds to variant n° 9

> datavar9 <- dataset[dataset$Variant == 9, -1]

> print(head(datavar9))

ID Y A B Visit

19579 1 54.16743 1 1 0

19580 1 33.62468 1 1 1

19581 1 16.39175 1 2 2

19582 1 40.94266 1 2 3

19583 1 31.66774 1 2 4

19584 2 41.43027 0 1 0

> print("Dimensions dataset - Variant 9")

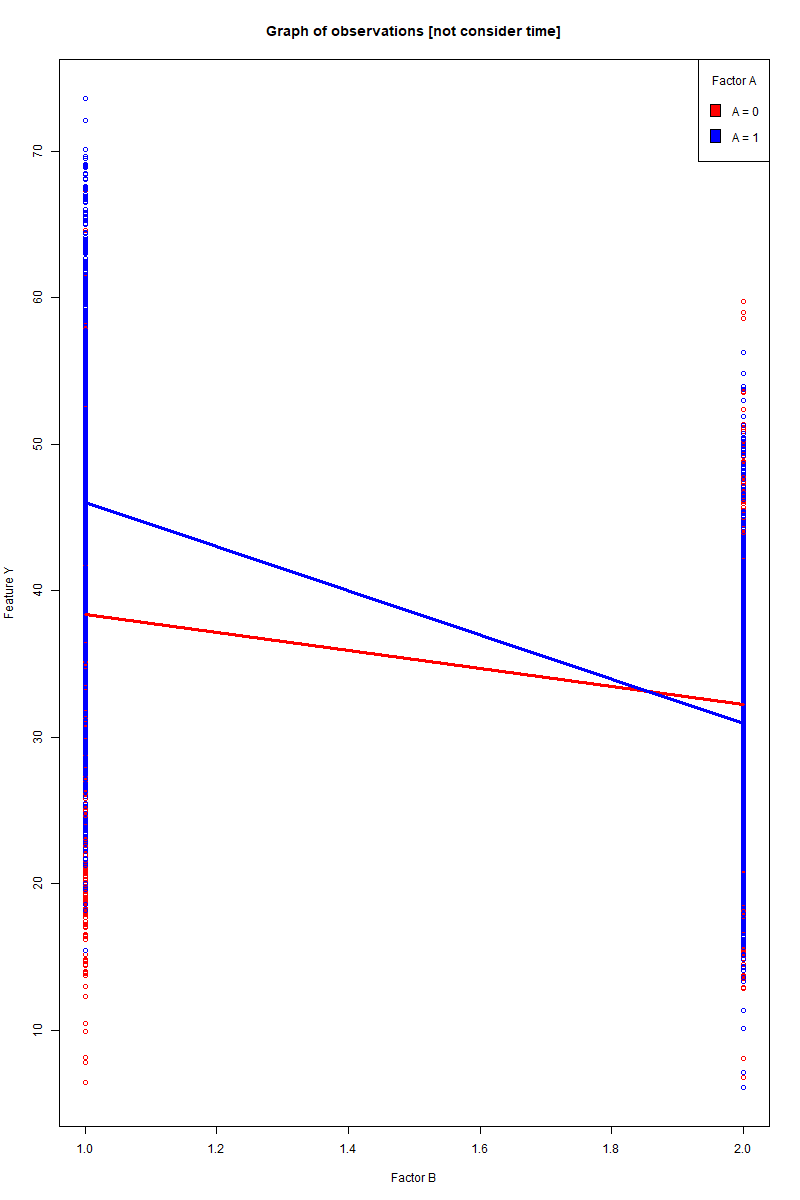
[1] "Dimensions dataset - Variant 9"

> print(dim(datavar9))

[1] 4405 5

**2. Изобразить графически результаты наблюдений без учета времени измерений исследуемой характеристики. В предположении нормальности значений наблюдаемого признака провести двухфакторный дисперсионный анализ зависимости наблюдаемого признака от значений факторов А и В.**

А) График построенa:



В) Двухфакторный дисперсионный анализ

Шаг 01: Модель взаймодействия между факторов А и В



> qab <- lm(Y ~ as.factor(A) \* as.factor(B), data = datavar9)

> qab

Call:

lm(formula = Y ~ as.factor(A) \* as.factor(B), data = datavar9)

Coefficients:

(Intercept) as.factor(A)1

38.354 7.679

as.factor(B)2 as.factor(A)1:as.factor(B)2

-6.149 -8.971

Шаг 02: Аддитивные модели

- Фактроры А и В:



> qadd <- lm(Y ~ as.factor(A) + as.factor(B), data = datavar9)

> qadd

Call:

lm(formula = Y ~ as.factor(A) + as.factor(B), data = datavar9)

Coefficients:

(Intercept) as.factor(A)1 as.factor(B)2

40.49 3.11 -10.33

- Только Фактор А:



> qa <- lm(Y ~ as.factor(A), data = datavar9)

> qa

Call:

lm(formula = Y ~ as.factor(A), data = datavar9)

Coefficients:

(Intercept) as.factor(A)1

35.217 3.128

- Только фактор В:



> qb <- lm(Y ~ as.factor(B), data = datavar9)

> qb

Call:

lm(formula = Y ~ as.factor(B), data = datavar9)

Coefficients:

(Intercept) as.factor(B)2

41.94 -10.34

- Независимые от факторов А и В:



> q0 <- lm(Y ~ 1, data = datavar9)

> q0

Call:

lm(formula = Y ~ 1, data = datavar9)

Coefficients:

(Intercept)

36.68

Шаг 03: Проверка гипотез для двухфакторный дисперсионный анализ

Пусть модель: 

А) Гипотеза отсутствия взаймодействий



> anova(qadd, qab)

Analysis of Variance Table

Model 1: Y ~ as.factor(A) + as.factor(B)

Model 2: Y ~ as.factor(A) \* as.factor(B)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 4402 401345

2 4401 379295 1 22050 255.85 < 2.2e-16 \*\*\*

---

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

Гипотеза  **отвергает**, следовательно, есть влияние взаимодействия между А и В.

Б) Гипотеза отсутствие влияния фактора А



> anova(qa, qab)

Analysis of Variance Table

Model 1: Y ~ as.factor(A)

Model 2: Y ~ as.factor(A) \* as.factor(B)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 4403 518919

2 4401 379295 2 139624 810.03 < 2.2e-16 \*\*\*

---

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

Гипотезы  **и отвергают**, следовательно, есть аддитивный влияние фактора А.

В) Гипотеза отсутствие влияния фактора В



> anova(qb, qab)

Analysis of Variance Table

Model 1: Y ~ as.factor(B)

Model 2: Y ~ as.factor(A) \* as.factor(B)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 4403 411951

2 4401 379295 2 32656 189.45 < 2.2e-16 \*\*\*

---

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

Гипотезы  **и отвергают**, следовательно, есть аддитивный влияние фактора B.

Г) Гипотеза независимости Y от факторов А, В.



> anova(q0, qab)

Analysis of Variance Table

Model 1: Y ~ 1

Model 2: Y ~ as.factor(A) \* as.factor(B)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 4404 529645

2 4401 379295 3 150350 581.51 < 2.2e-16 \*\*\*

---

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

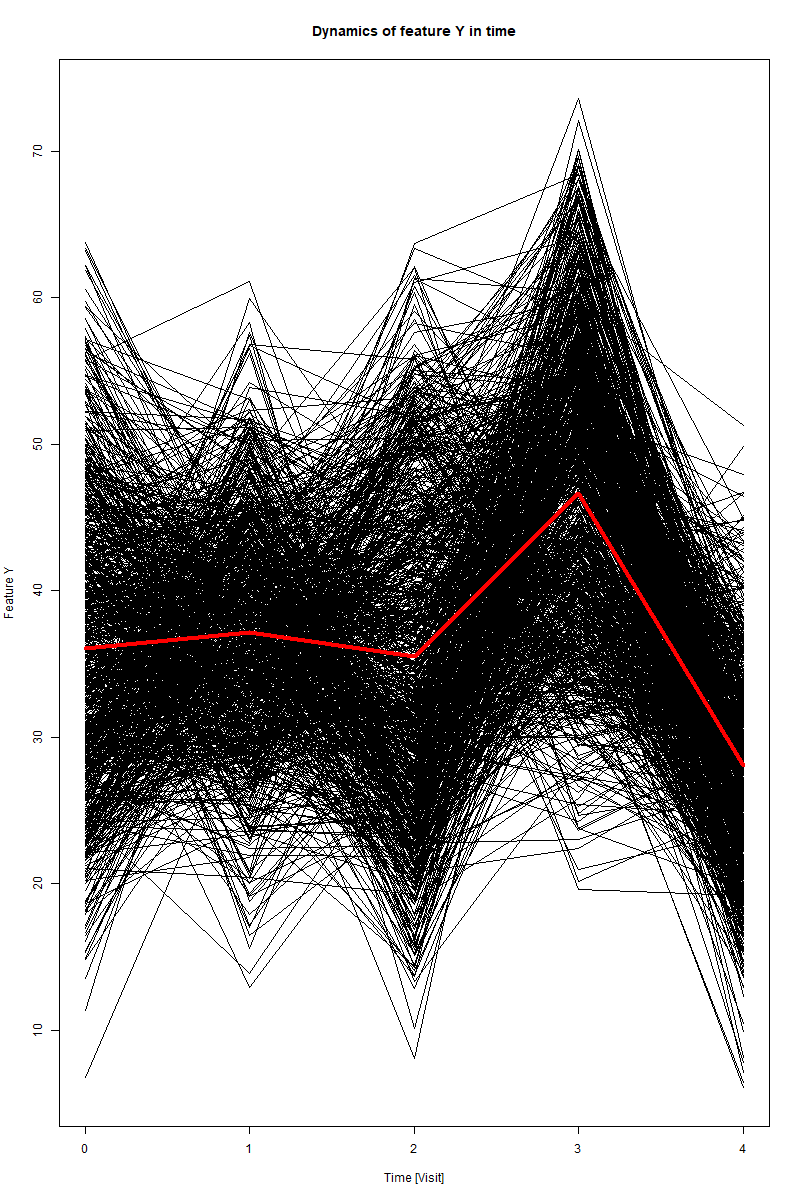
Гипотезы , и отвергают

Общий вывод:

Есть влиянгия взаймодейство и аддитивнное для факторов А и В с отношениями наблюдаемого признака.

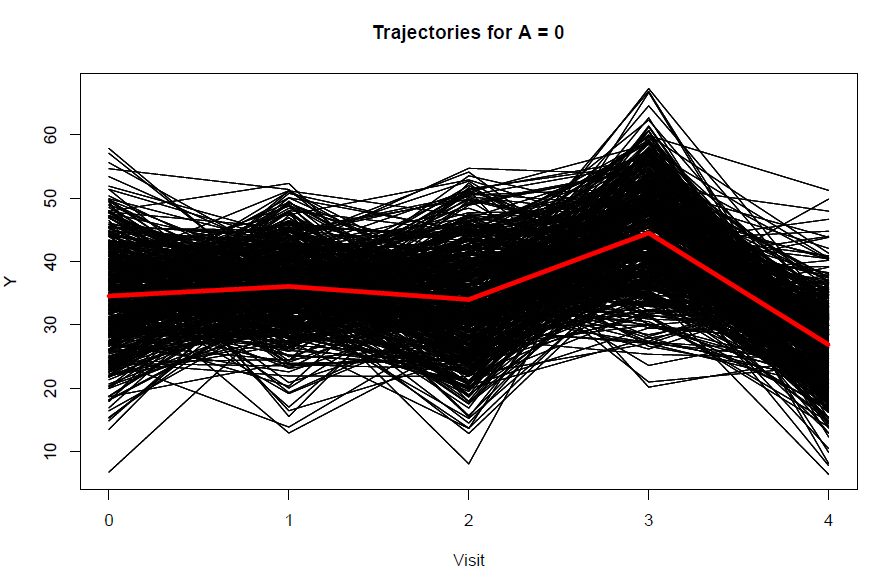
**3. Преставить визуально динамику изменений наблюдаемого признака в виде траекторий и оценить наличие влияния постоянного фактора А на значения измеряемого признака. Оценить корреляции значений наблюдаемого признака в различные моменты времени при каждом значений признака А без учета влияния признака В.**

А) График построенa динамику изменений наблюдаемого признака

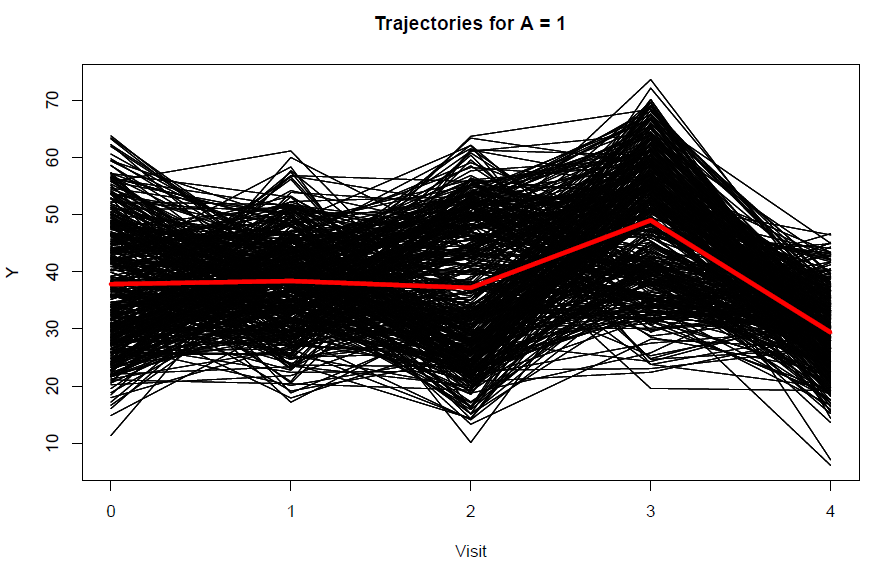


Б) Оценить наличие вляния постоянного фактора А на значения Y.

- Траектория динамику Y (во времени) для А = 0



- Траектория динамику Y (во времени) для А = 1



В) Оценить корреляции значения Y (во времени) при каждом значения А.

Выявляем возможные значения переменной А

> lvA = levels(as.factor(datavar9$A))

[1] "0" "1"

Корреляционная матрица для Y (во времени) при A = 0

> corr\_A0

$`A=0`$corr

Y\_V0 Y\_V1 Y\_V2 Y\_V3 Y\_V4

Y\_V0 1.00000000 -0.026984902 -0.036062662 -0.04480147 0.05599858

Y\_V1 -0.02698490 1.000000000 0.008961128 0.09223862 0.02129035

Y\_V2 -0.03606266 0.008961128 1.000000000 -0.07581508 0.05553220

Y\_V3 -0.04480147 0.092238617 -0.075815077 1.00000000 -0.05049627

Y\_V4 0.05599858 0.021290351 0.055532204 -0.05049627 1.00000000

Корреляционная матрица для Y (во времени) при A = 1

> corr\_A1

$`A=1`$corr

Y\_V0 Y\_V1 Y\_V2 Y\_V3 Y\_V4

Y\_V0 1.00000000 0.017432803 0.01894922 0.034281278 -0.02109358

Y\_V1 0.01743280 1.000000000 -0.04929156 -0.001735304 0.06703278

Y\_V2 0.01894922 -0.049291562 1.00000000 -0.039914042 0.02080742

Y\_V3 0.03428128 -0.001735304 -0.03991404 1.000000000 -0.06104107

Y\_V4 -0.02109358 0.067032779 0.02080742 -0.061041066 1.00000000

**4. Построит оценки средних значений наблюдаемой характеристики при различных значенях факторов А и В и центрировать исходные наблюдения. С использованием семивариограммы, оценить зависимость корреляции значений центрированного процесса от времени в предположении его стационарности при каждом значении пары факторов А и В.**

Шаг 01: Для оценить средних значений наблюдаемой характеристики при различных значениях факторов А и В, получиться смещанную линейную модель по каждому пары значений факторов А и В.

> # generate the linear mixed effect model for each pair values A, B

> m0 <- NULL

> for(i in names(d9)){

m0[[i]] <- lme(Y ~ as.factor(Visit), random = ~1|ID, data = d9[[i]])

}

Шаг 02: Построить семивариограммы по каждому пары значений факторов А и В.

Таблички:

> # check the table of semi-variogram distances and variogram values

$`a=0\_b=1` $`a=0\_b=2`

variog dist n.pairs variog dist n.pairs

1 1.0975740 1 445 1 1.0469750 1 504

2 0.9219419 2 330 2 0.9429442 2 367

3 0.9705741 3 232 3 1.0097347 3 243

4 1.0422249 4 121 4 0.9548429 4 110

$`a=1\_b=1` $`a=1\_b=2`

variog dist n.pairs variog dist n.pairs

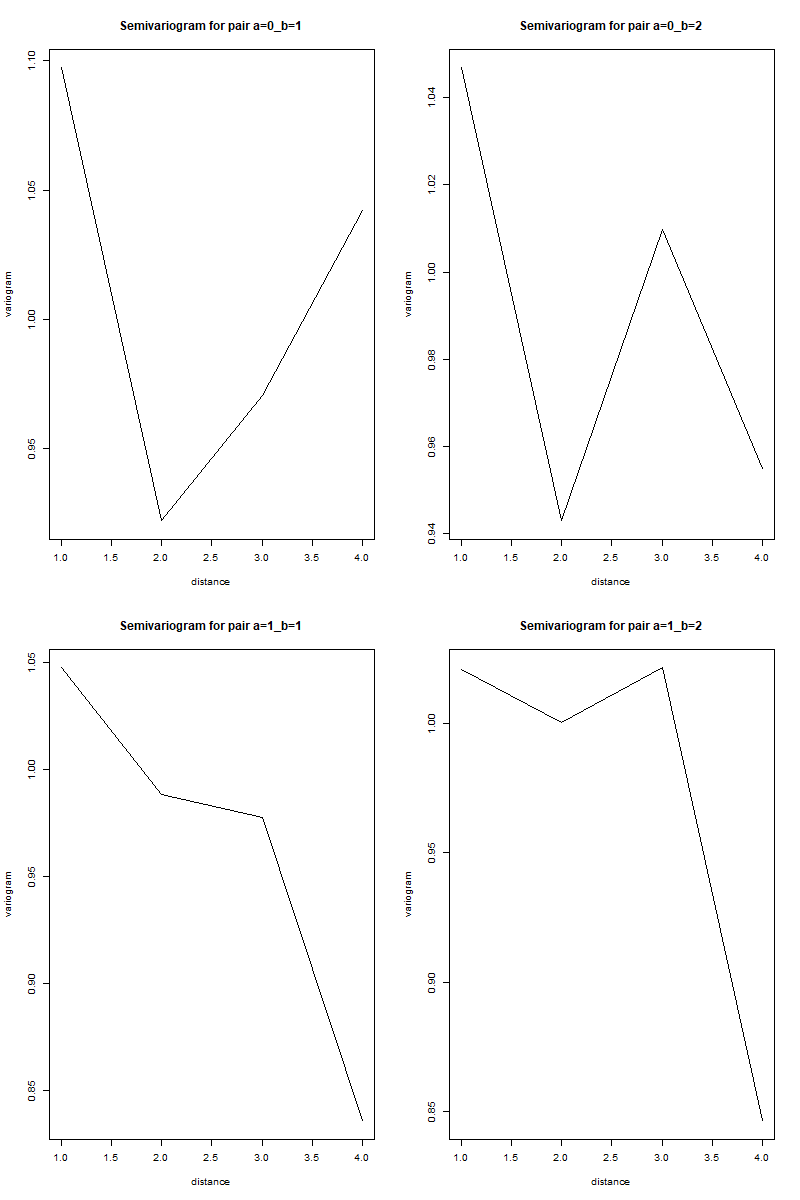
1 1.0479601 1 398 1 1.0206838 1 419

2 0.9884016 2 304 2 1.0002901 2 336

3 0.9778377 3 210 3 1.0216352 3 213

4 0.8359854 4 97 4 0.8464707 4 111

Графики семивариограммы:



Шаг 03: Оценить зависимость корреляции значений центрированного процесса от времени в предположении его стационарности при каждом значении пары факторов А и В

> print(corr)

$`a=0\_b=1`$corr

[Visit = 0] [Visit = 1] [Visit = 2] [Visit = 3] [Visit = 4]

[1,] 1.00000000 -0.09757398 0.07805808 0.02942590 -0.04222494

[2,] -0.09757398 1.00000000 -0.09757398 0.07805808 0.02942590

[3,] 0.07805808 -0.09757398 1.00000000 -0.09757398 0.07805808

[4,] 0.02942590 0.07805808 -0.09757398 1.00000000 -0.09757398

[5,] -0.04222494 0.02942590 0.07805808 -0.09757398 1.00000000

$`a=0\_b=2`$corr

[Visit = 0] [Visit = 1] [Visit = 2] [Visit = 3] [Visit = 4]

[1,] 1.000000000 -0.046974985 0.05705581 -0.009734725 0.045157074

[2,] -0.046974985 1.000000000 -0.04697499 0.057055806 -0.009734725

[3,] 0.057055806 -0.046974985 1.00000000 -0.046974985 0.057055806

[4,] -0.009734725 0.057055806 -0.04697499 1.000000000 -0.046974985

[5,] 0.045157074 -0.009734725 0.05705581 -0.046974985 1.000000000

$`a=1\_b=1`$corr

[Visit = 0] [Visit = 1] [Visit = 2] [Visit = 3] [Visit = 4]

[1,] 1.00000000 -0.04796008 0.01159842 0.02216233 0.16401459

[2,] -0.04796008 1.00000000 -0.04796008 0.01159842 0.02216233

[3,] 0.01159842 -0.04796008 1.00000000 -0.04796008 0.01159842

[4,] 0.02216233 0.01159842 -0.04796008 1.00000000 -0.04796008

[5,] 0.16401459 0.02216233 0.01159842 -0.04796008 1.00000000

$`a=1\_b=2`$corr

[Visit = 0] [Visit = 1] [Visit = 2] [Visit = 3] [Visit = 4]

[1,] 1.00000000 -0.02068379 -0.00029010 -0.02163522 0.15352928

[2,] -0.02068379 1.00000000 -0.02068379 -0.00029010 -0.02163522

[3,] -0.00029010 -0.02068379 1.00000000 -0.02068379 -0.00029010

[4,] -0.02163522 -0.00029010 -0.02068379 1.00000000 -0.02068379

[5,] 0.15352928 -0.02163522 -0.00029010 -0.02068379 1.00000000

**5. Пострить смещанную линейную с простым эффектом индивида, а также смещанную линейную модель, допускающую линейную зависимость эффекта индивида от времени.**

Шаг 01: Построение смешанную линейную модель с простым эффектом индивида.

> # Build linear mixed effect model (lme) with simple individual effect

> q51 <- lme(fixed = Y ~ as.factor(Visit), random =~ 1|ID,

+ data = datavar9, method = "ML")

> print(q51)

Linear mixed-effects model fit by maximum likelihood

Data: datavar9

Log-likelihood: -16039.02

Fixed: Y ~ as.factor(Visit)

(Intercept) as.factor(Visit)1 as.factor(Visit)2 as.factor(Visit)3

36.0884731 1.0633315 -0.5996778 10.5155339

as.factor(Visit)4

-8.0404377

Random effects:

Formula: ~1 | ID

(Intercept) Residual

StdDev: 1.447286 9.118381

Number of Observations: 4405

Number of Groups: 881

Шаг 02: Построение смешанную линейную модель с линейную зависимость эффекта индивида от времени

> # Linear mixed model with linear dependency between individual effect and time

> q52 <- lme(fixed = Y ~ as.factor(Visit), random =~ (1 + Visit)|ID,

+ data = datavar9, method = "ML")

> print(q52)

Linear mixed-effects model fit by maximum likelihood

Data: datavar9

Log-likelihood: -16039.02

Fixed: Y ~ as.factor(Visit)

(Intercept) as.factor(Visit)1 as.factor(Visit)2 as.factor(Visit)3

36.0884731 1.0633315 -0.5996778 10.5155339

as.factor(Visit)4

-8.0404377

Random effects:

Formula: ~(1 + Visit) | ID

Structure: General positive-definite, Log-Cholesky parametrization

StdDev Corr

(Intercept) 1.4472896795 (Intr)

Visit 0.0004759267 -0.001

Residual 9.1183808651

Number of Observations: 4405

Number of Groups: 881

Шаг 03: Чтобы определить если целесообразность введения коэффициента наклона случайного эффекта, мы пользуем критерия cAIC.

> # apply cAIC criteria to establish the best model

> tblcAIC <- data.frame(models = c(model1, model2),

+ cAICvalues = c(q51\_cAIC$caic, q52\_cAIC$caic))

>

> # table of cAIC values

> print(tblcAIC)

models cAICvalues

1 y ~ Visit + (1|ID) 32087.86

2 y ~ Visit + (1+Visit|ID) 32099.03

Модель 1 является лучшей моделью, так как она имеет более низкое значение индекса cAIC.

**6. С учетом результатов п.5, пострить смешанную модель зависимости наблюдаемого признака от значений признаков А и В с учетом времени наблюдения. При каждом значений фактора А проверить гипотезы аддитивности влияния фактора В и времени наблюдения а также гипотезы отсутствия влияния каждого из факторов В и времени наблюдения. Включить в модель фактор А и оценить вляние фактора А на значение наблюдаемой характеристики.**

Шаг 01: Смешанная модель зависимости А, В и времени наблюдения с простым эффектом индивида: 

> # mixed model: (Y|A, B, Visit)

> # model: nijk = u + aA\_i + aB\_j + aT\_k +

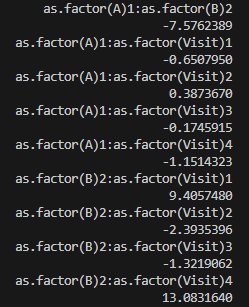
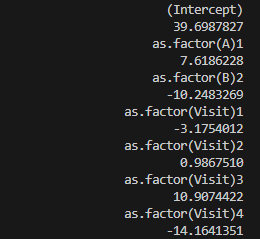
> # aA\_i \* aB\_j + aA\_i \* aT\_k + aB\_j \* aT\_k + aABT\_ijk

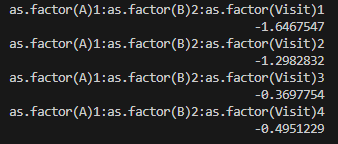
> q61 <- lme(fixed = Y ~ as.factor(A) \* as.factor(B) \* as.factor(Visit),

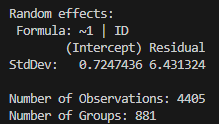
+ random =~ 1|ID, data = datavar9, method = "ML")

Linear mixed-effects model fit by maximum likelihood

Data: datavar9

 Fixed: Y ~ as.factor(A) \* as.factor(B) \* as.factor(Visit)

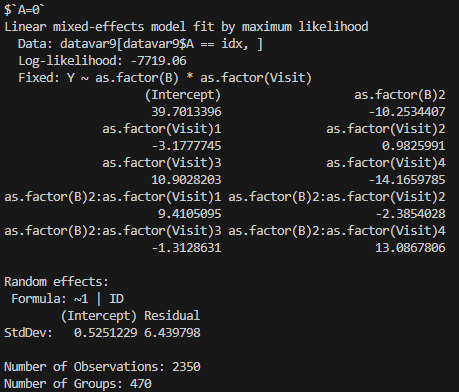


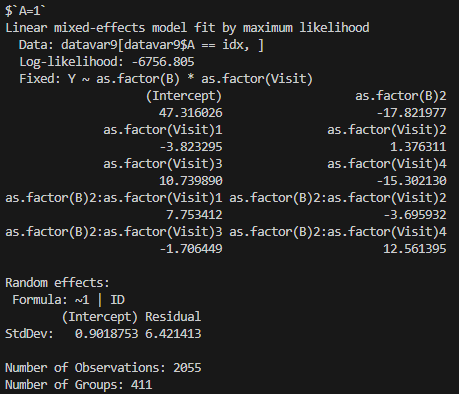


Шаг 02: Модель для взаймодействия между фактора В и времени наблюдения по каждому значения факторa А: 

> # interaction model for factor B and time (for each value of A)

> # model: nij = u + aB\_j + aT\_k + aBT\_jk

> q6\_intBT

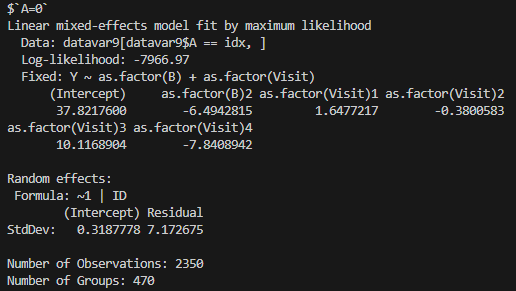


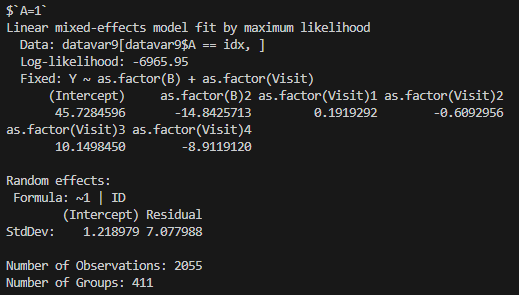
Шаг 03: Аддитивный модель для фактора В и времени наблюдения по каждому значения факторa А: 

> # additive model for factor B and time (for each value of A)

> # model: njk = u + aB\_j + aT\_k

> q6\_addBT



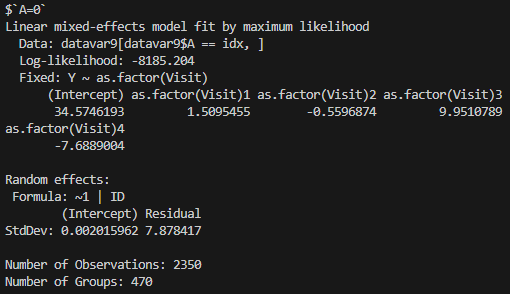


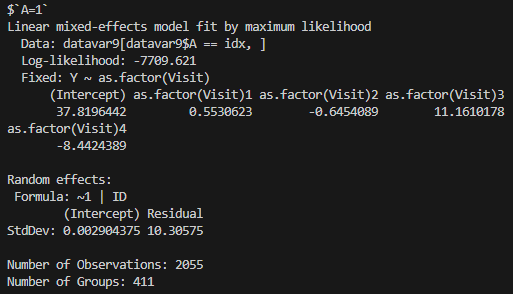
Шаг 04: Аддитивный модель только для фактора времени по каждому значения факторa А: 

> # additive model for time (for each value of A)

> # model: njk = u + aT\_k

> q6\_addT

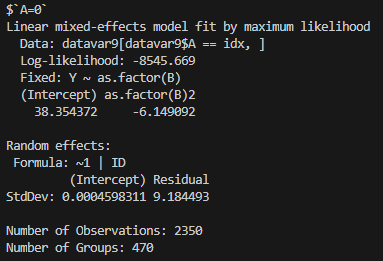


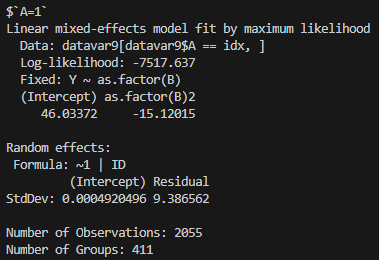


Шаг 05: Аддитивный модель только для фактора В по каждому значения факторa А: 

> # additive model for factor B (for each value of A)

> # model: njk = u + aB\_j

> q6\_addB



Шаг 06: Проверка гипотезы

H\_BT: Гипотеза аддитивности влияния фактора В и времени наблюдения



H\_B: Гипотеза отсутствие влияния фактора В



H\_T: Гипотеза отсутствие влияния фактора времени наблюдения



> # check the p-values to make the hypothesis inference

> print(tbl\_hipothesis())

A = 0 A = 1

H\_BT: bBT\_ij = 0 5.372022e-106 3.106080e-89

H\_B: bBT\_ij = 0 and bB\_i = 0 2.733439e-199 0.000000e+00

H\_T: bBT\_ij = 0 and bT\_j = 0 0.000000e+00 2.964394e-323

Шаг 07: Заключение

Для А = 0 и А = 1:

Гипотеза **отвергает**, тогда есть влияние взаймодействия между фактора В и времени наблюдения

Гипотезы  **и отвергают**, тогда есть аддитивности влияния фактора В

Гипотезы  **и**  **отвергают**, тогда есть аддитивности влияния фактора времени наблюдения

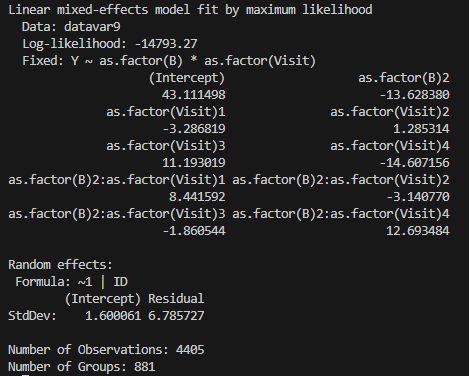
Шаг 08: Построить модель с отсутсвием влияния фактора А



> # model without factor A

> q62 <- lme(fixed = Y ~ as.factor(B) \* as.factor(Visit),

+ random =~ 1|ID, data = datavar9, method = "ML")



Шаг 09: Оценить вляние фактора А на значение наблюдаемой характеристика Y

Пусть модели:

Модель 1: 

Модель 2: 

Гипотеза HA: 

> # study the influence of factor A

> anov\_addA <- anova(q62, q61)

> print(anov\_addA)

Model df AIC BIC logLik Test L.Ratio p-value

q62 1 12 29610.53 29687.22 -14793.27

q61 2 22 28996.08 29136.67 -14476.04 1 vs 2 634.452 <.0001

> print(paste("p-value = ", anov\_addA$"p-value"[2]))

[1] "p-value = 7.26550582148688e-130"

Шаг 10: Заключение

Гипотеза HA **отвергает**, тогда есть влияние фактора А на значение наблюдаемой характеристики.

> # check cAIC values

> q62\_aic <- cAIC(q62)$caic

> print(q62\_aic)

[1] 29586.8

> q61\_aic <- cAIC(q61)$caic

> print(q61\_aic)

[1] 28996.23

С критерием AIC, тоже видем что модель с влияния фактора А – это лучшее модель чем модель с отсутствием влияния фактора А.

**8. Построить смешанную модель ковариационного анализа в предположении полиномиальной зависимости второго порядка наблюодаемого признака от времени. Проверить гипотезу линейности зависимости средного значения наблюдаемого признака от времени в пристствии факторов А и В.**

Шаг 01: Пострить смешанную модель для полиномиальной зависимости второго порядка наблюдаемого признака от времени



> # build model mixed-analysis for 2nd polynomial order in time

> # model: nijkl = u + aA\_i + aB\_j + aT\_k \* T + aT^2\_l \*T^2 +

> # aAB\_ij + aAT\_ik \* T + aBT\_jk \* T + aAT^2\_il \* T^2 + aBT^2\_jl \* T^2 +

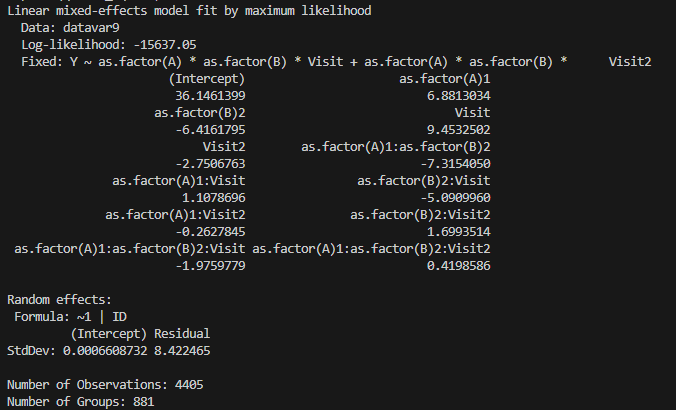
> # aABT\_ijk \* T + aABT^2\_ijl \* T^2

> q8.abt\_square <- lme(fixed = Y ~ as.factor(A) \* as.factor(B) \* Visit +

+ as.factor(A) \* as.factor(B) \* Visit2,

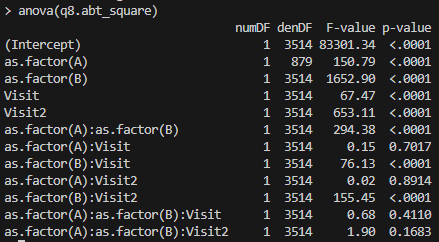
+ random =~ 1|ID, data = datavar9, method = "ML")

> print(q8.abt\_square)



> # check the significance for each coefficient

> anova(q8.abt\_square)



Шаг 02: Пострить смешанную модель с линейности зависимости средного значения наблдаемого признака от времени в присутствии факторов А и В



> # build model mixed-analysis for linear dependency

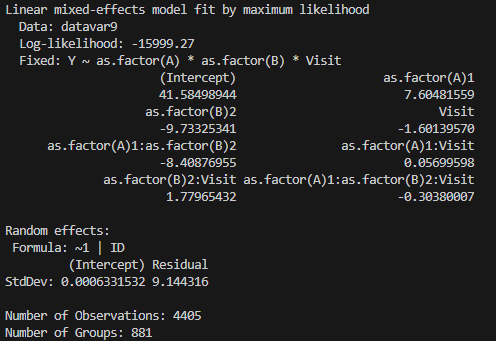
> # model: nijk = u + aA\_i + aB\_j + aT\_j \* T + aAB\_ij + aAT\_ik \* T + aBT\_jk +

aABT\_ijk \* T

> q8.abt\_linear <- lme(fixed = Y ~ as.factor(A) \* as.factor(B) \* Visit,

+ random =~ 1|ID, data = datavar9, method = "ML")

> print(q8.abt\_linear)



Шаг 03: Проверка гипотеза линейности зависимости



> # proof of hypothesis for linear dependency

> # H\_T^2: aABT^2\_ijl = 0 and aAT^2\_il = 0 and aBT^2\_jl = 0 and aT^2\_l = 0

> q8\_anova1 <- anova.lme(q8.abt\_linear, q8.abt\_square)

> print(q8\_anova1)

Model df AIC BIC logLik Test L.Ratio p-value

q8.abt\_linear 1 10 32018.55 32082.45 -15999.27

q8.abt\_square 2 14 31302.10 31391.57 -15637.05 1 vs 2 724.4463 <.0001

>

> # check the p-value for H\_T^2

> q8\_pvalue1 <- q8\_anova1$"p-value"[2]

> print(q8\_pvalue1)

[1] 1.772828e-155

Шаг 04: Заключение:

Гипотеза отвергает, тогда есть влияние по полиномиальной зависимости второго порядка.

> # check the AIC values

> q8square\_caic <- cAIC(q8.abt\_square)$caic

> print(q8square\_caic)

[1] 31308.41

> q8linear\_caic <- cAIC(q8.abt\_linear)$caic

> print(q8linear\_caic)

[1] 32029.95

С пользованием условного информационного критерия Акайке (сAIC), видем что модель второго порядка для времени наблюдения – это лучшее чем просто линейной зависимости

**9. С использованием информационных критериев AIC и BIC, выбрать найлучшую модель для неслучайного эффекта в рамках смешанной модели с простым эффектом индивида. Провести исследование влияния факторов А, В и времени наблюдения на значение исследуемой характеристики в рамках данной модели.**

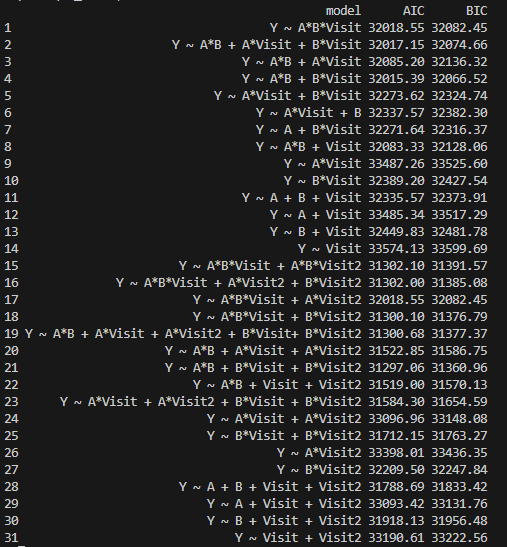
**Сравнить результаты, полученные с использованием AIC, и выбор найлучшей модели с учетом случайнного эффекта при помощи cAIC.**

Шаг 01: Проверяем значения информационных критериев AIC и BIC для каждой рассматриваемой модели.

> # build the information-criteria table which contains AIC and BIC values

> IC\_table <- data.frame(model = func, AIC = aic\_values, BIC = bic\_values)

> print(IC\_table)



Шаг 02: Выбираем лучшую модель по критериям AIC и BIC.

> # take the best model for AIC criteria

> best\_modelAIC <- IC\_table[IC\_table$AIC == min(IC\_table$AIC), ]

> print(best\_modelAIC)

id model AIC BIC

21 Y ~ A\*B + B\*Visit + B\*Visit2 31297.06 31360.96

> # take the best model for BIC criteria

> best\_modelBIC <- IC\_table[IC\_table$BIC == min(IC\_table$BIC), ]

> print(best\_modelBIC)

id model AIC BIC

21 Y ~ A\*B + B\*Visit + B\*Visit2 31297.06 31360.96

Шаг 03: Провести исследование влияния факторов А, В и времени наблюдения на значение времени наблюдаемой характеристики в самой лучшей модели.

Самое лучшей модель: Y ~ A\*B + B\*Visit + B\*Visit2



> # check the parameters inside of the optimal model

> opt\_id <- IC\_table[IC\_table$AIC == min(IC\_table$AIC), ]$id

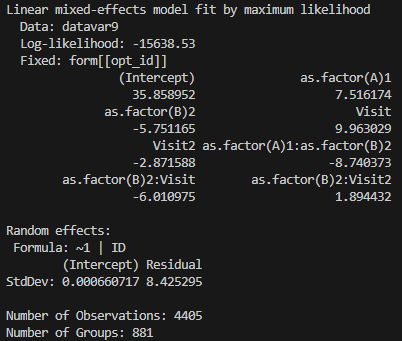
> print(paste("opt-id = ", opt\_id))

[1] "opt-id = 21"

> model\_opt <- lme(fixed = form[[opt\_id]],

+ random =~ 1|ID, data = datavar9, method = "ML")

> print(model\_opt)

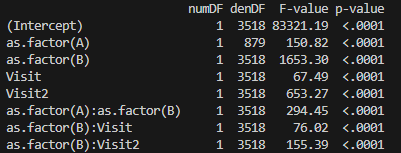


Значимость параметров оптимальной модели:

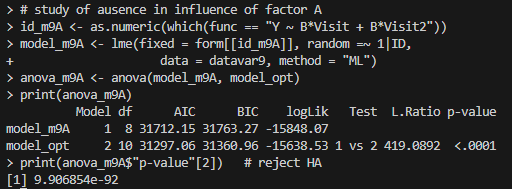
> # check the significance of the parameters in the optimal model

> anova\_opt <- anova(model\_opt)

> print(anova\_opt)

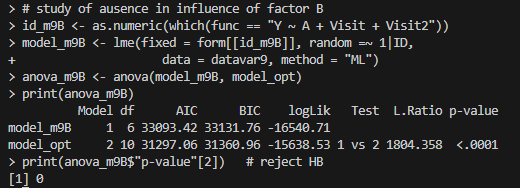


- отсутствие влияния фактор А: 



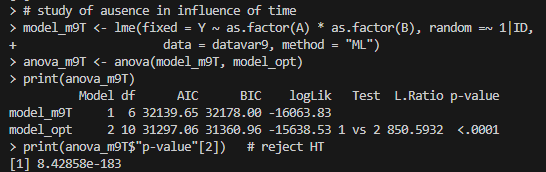
Отвергает HA, тогда есть влияния фактора А в найлучшей модели.

- отсутствие влияния фактор В: 



Отвергает HB, тогда есть влияния фактора В в найлучшей модели.

- отсутствие влияния фактор времени наблюдения: 



Отвергает HT, тогда есть влияния фактора врмени в найлучшей модели.

Шаг 04: Сравнить результаты и выбор найлучшей модели с учетом случайнного эффеута при помощи cAIC.

Модель: Y ~ A\*B + B\*Visit + B\*Visit2

> # compare results of AIC and cAIC for best model

> aic\_best <- AIC(model\_opt)

> print(aic\_best)

[1] 31297.06

> caic\_best <- cAIC(model\_opt)

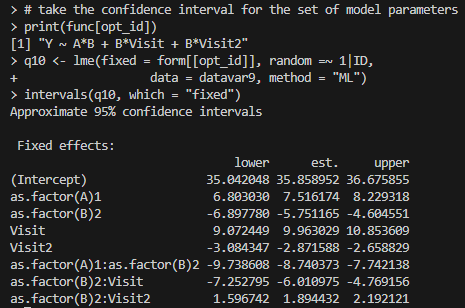
> print(caic\_best$caic)

[1] 31303.36

**10. Посчитать частные и совместные доверительные интервалы (прнимает** с **простой эффект индивида).**

Шаг 01: Частное доверительные доверительные интервалы: 

> intervals(q10, which = "fixed", level = 1 - alpha)



Шаг 02: Совместное доверительные доверительные интервалы

- использовать метод Bonferroni: 

> alpha2 <- alpha / length(q10$coefficients[[1]])

> print(paste("alpha for bonferroni = ", alpha2))

[1] "alpha for bonferonii = 0.00625"

> intervals(q10, which = "fixed", level = 1 - alpha2)

