

Смешанные линейные модели (вложенные случайные факторы)

Линейные модели...

Марина Варфоломеева, Вадим Хайтов

СПбГУ



Вы узнаете

- ▶ Что такое вложенные случайные факторы и в каких случаях они применяются

Вы сможете

- ▶ Объяснить, что такое вложенные случайные факторы
- ▶ Привести примеры иерархических случайных факторов
- ▶ Вычислить коэффициент внутриклассовой корреляции для случая с двумя вложенными случайными факторами
- ▶ Подобрать смешанную линейную модель со вложенными случайными факторами

Смешанные модели со вложенными случайными факторами

Вложенные факторы (Nested effects)



Вложенные факторы (Nested effects)

Факторы образуют иерархическую последовательность вложенности

- ▶ лес → дерево в лесу → ветка на дереве → наблюдение (личинки насекомых)

Вложенные факторы (Nested effects)

Факторы образуют иерархическую последовательность вложенности

- ▶ лес → дерево в лесу → ветка на дереве → наблюдение (личинки насекомых)

Внутри каждого уровня главного фактора будут разные (нестрого сопоставимые) уровни вложенного фактора

Деревья, с которых собирали личинок, будут разные в разных лесах (разные экземпляры).

Вложенные факторы (Nested effects)

Факторы образуют иерархическую последовательность вложенности

- ▶ лес → дерево в лесу → ветка на дереве → наблюдение (личинки насекомых)

Внутри каждого уровня главного фактора будут разные (нестрого сопоставимые) уровни вложенного фактора

Деревья, с которых собирали личинок, будут разные в разных лесах (разные экземпляры).

Уровни вложенных факторов описывают иерархию взаимного сходства наблюдений

Личинки с разных деревьев из одного леса имеют право быть похожими друг на друга больше, чем на личинок из другого леса

Личинки на одном дереве имеют право быть похожими друг на друга больше, чем на личинок с другого дерева

И т.п.

Другие примеры вложенных факторов

- ▶ регион -> город -> больница -> наблюдение (пациент)
- ▶ самка -> выводок -> наблюдение (особь)
- ▶ лес -> дерево в лесу -> гнездо на дереве -> наблюдение (птенец)
- ▶ улитка -> спороциста в улитке -> наблюдение (редия)

Пример: Высота растений и выпас скота

Вообще-то, статья Gennet et al. 2017 о птицах, но чтобы про них что-то лучше понять, нужно разобраться с их местообитанием.

Как в разные годы высота растительного покрова зависит от выпаса скота, экспозиции склона и проективного покрытия местных растений?

Зависимая переменная:

- ▶ **height** - высота растительного покрова

Предикторы:

- ▶ **graze** - выпас коров (0, 1)
- ▶ **AspectCat** - экспозиция (S, N)
- ▶ **nativecov** - покрытие местной флоры %
- ▶ **slope** - наклон
- ▶ **year** - год наблюдений
- ▶ **Park** - парк
- ▶ **plotID** - уникальный идентификатор участка

Данные: Gennet, S., Spotswood, E., Hammond, M. and Bartolome, J.W., 2017. Livestock grazing supports native plants and songbirds in a California annual grassland. PloS one, 12(6), p.e0176367.



Открываем данные

Откроем и переформатируем данные так, чтобы не было дублирования и каждому участку соответствовала одна строчка.

```
library(readxl)
library(tidyr)
gr <- read_excel("data/Grazing_native_plants_Gennet_et_al._2017_S1.xlsx")
graz <- gr %>% spread(Species, presence)
```

Знакомство с данными

Есть ли пропущенные значения?

```
sum(is.na(graz))
```

```
# [1] 0
```

Сколько участков было в каждом парке в каждый год?

```
with(graz, table(Park, year))
```

```
#      year
# Park 2004 2005 2006 2007 2008 2009 2010 2011
#  MT     6   10   10   10   10   10   10   10
#  PR     6    6    6    6    6    6    6    6
#  SU     0    9    9    9    9    9    9    9
#  VC    10   10   10   10   11   11   11   11
```

Наводим порядок

Сделаем факторами переменные, которые понадобятся для модели

```
graz$graze_f <- factor(graz$graze)
graz$AspectCat <- factor(graz$AspectCat)
graz$year_f <- factor(graz$year)
```

Извлечем корень из обилия местных видов

```
graz$nativecov_sq <- sqrt(graz$nativecov)
```

Вспомним главный вопрос исследования и подберем модель

Как в разные годы высота растительного покрова зависит от выпаса скота, экспозиции склона и проективного покрытия местных растений?

Вспомним главный вопрос исследования и подберем модель

Как в разные годы высота растительного покрова зависит от выпаса скота, экспозиции склона и проективного покрытия местных растений?

Нам нужно учесть, что в разные годы из-за кучи разных причин высота растений может различаться

Кроме того, нужно учесть, что в разных парках и на разных участках растения будут расти сходным образом в разные годы. У нас есть иерархические факторы парк и участок в парке

Модель

Вспомним главный вопрос исследования и подберем модель

Как в разные годы высота растительного покрова зависит от выпаса скота, экспозиции склона и проективного покрытия местных растений?

Нам нужно учесть, что в разные годы из-за кучи разных причин высота растений может различаться

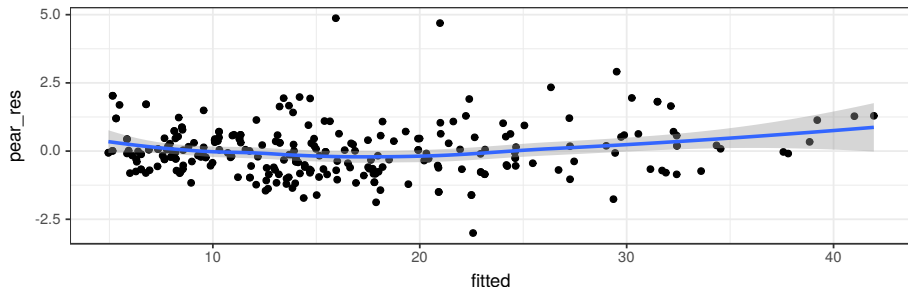
Кроме того, нужно учесть, что в разных парках и на разных участках растения будут расти сходным образом в разные годы. У нас есть иерархические факторы парк и участок в парке

```
library(nlme)
MN1 <- lme(height ~ graze_f*AspectCat + year_f + nativecov_sq + slope,
           random = ~ 1|Park/plotID,
           data = graz, method = "ML")
```

```
# Данные для анализа остатков  
MN1_diag <- data.frame(  
  graz,  
  pear_res = residuals(MN1, type = "pearson"),  
  fitted = fitted(MN1, type = "response"))
```

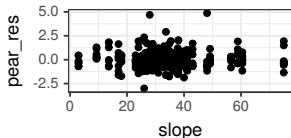
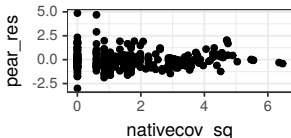
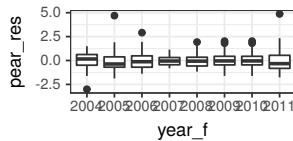
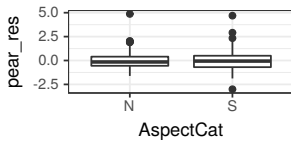
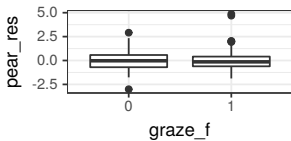

График остатков

```
library(ggplot2)
gg_res <- ggplot(data = MN1_diag, aes(y = pear_res))
gg_res + geom_point(aes(x = fitted)) +
  geom_smooth(aes(x = fitted))
```



Графики остатков от переменных в модели

```
library(gridExtra)
grid.arrange(gg_res + geom_boxplot(aes(x = graze_f)),
gg_res + geom_boxplot(aes(x = AspectCat)),
gg_res + geom_boxplot(aes(x = year_f)),
gg_res + geom_point(aes(x = nativecov_sq)),
gg_res + geom_point(aes(x = slope)),
ncol = 3)
```

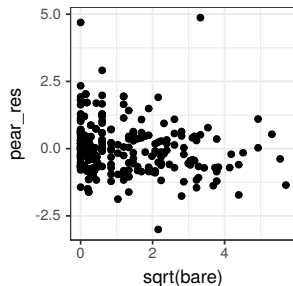
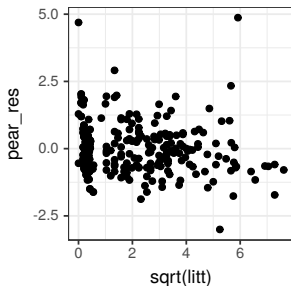
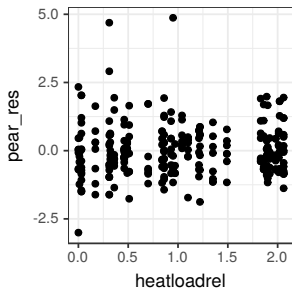


- Паттерн на графике nativecov_sq. Возможно, здесь нужно использовать GAMM.



Графики остатков от переменных не в модели

```
grid.arrange(  
  gg_res + geom_point(aes(x = heatloadrel)),  
  gg_res + geom_point(aes(x = sqrt(litt))),  
  gg_res + geom_point(aes(x = sqrt(bare))),  
  ncol = 3)
```



Паттерн на графике heatloadrel - Возможно, есть тренд на графике sqrt(litt)

Тесты отношения правдоподобий для полной модели

Модель MN1 была подобрана при помощи ML, поэтому можно применять тесты отношения правдоподобий прямо к ней

```
library(car)  
Anova(MN1)
```

```
# Analysis of Deviance Table (Type II tests)  
#  
# Response: height  
#  
#               Chisq Df Pr(>Chisq)  
# graze_f        66.8805  1  2.885e-16 ***  
# AspectCat       1.8787  1  0.1704852  
# year_f        130.9568  7  < 2.2e-16 ***  
# nativecov_sq    2.3403  1  0.1260658  
# slope           0.5090  1  0.4755690  
# graze_f:AspectCat 11.5895  1  0.0006632 ***  
# ---  
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Высота растительного покрова:

- ▶ на склонах разной экспозиции по-разному зависит от выпаса скота (достоверное взаимодействие)
- ▶ различается в разные годы

Задание 3

Рассчитайте внутриклассовую корреляцию

- ▶ Для наблюдений на одном и том же участке
- ▶ Для наблюдений в одном и том же парке

Внутриклассовая корреляция

Для расчета внутриклассовой корреляции нужна модель, подобранная при помощи REML

```
MN1_fin <- lme(height ~ graze_f*AspectCat + year_f + nativecov_sq + slope,  
  random = ~ 1|Park/plotID,  
  data = graz, method = "REML")
```

MN1_fin

```
# Linear mixed-effects model fit by REML  
# Data: graz  
# Log-restricted-likelihood: -838.9555  
# Fixed: height ~ graze_f * AspectCat + year_f + nativecov_sq + slope  
# (Intercept) graze_f1 AspectCat5  
# 18.44908400 -5.25432888 8.57164897  
# year_f2005 year_f2006 year_f2007  
# 6.80750075 4.07972009 -0.58734178  
# year_f2008 year_f2009 year_f2010  
# -2.70376064 -2.69191822 -2.69191822  
# year_f2011 nativecov_sq slope  
# 3.72904015 -0.53116130 -0.02821586  
# graze_f1:AspectCat5  
# -9.44492743  
#  
# Random effects:  
# Formula: ~1 | Park  
# (Intercept)  
# StdDev: 1.574143  
#  
# Formula: ~1 | plotID %in% Park  
# (Intercept) Residual  
# StdDev: 3.3702 5.133291  
#  
# Number of Observations: 271  
# Number of Groups:  
# Park plotID %in% Park  
# 4 36
```



Внутриклассовая корреляция

Для наблюдений на одном и том же участке $\sigma_{plotID}^2 / (\sigma_{plotID}^2 + \sigma_{Park}^2 + \sigma^2)$

```
3.3702^2 / (1.574143^2 + 3.3702^2 + 5.133291^2)
```

```
# [1] 0.2826359
```

Для наблюдений в одном и том же парке $\sigma_{Park}^2 / (\sigma_{plotID}^2 + \sigma_{Park}^2 + \sigma^2)$

```
1.574143^2 / (1.574143^2 + 3.3702^2 + 5.133291^2)
```

```
# [1] 0.06166012
```

В результатах

Random effects:

Formula: ~1 | Park
(Intercept)

StdDev: 1.574143

Formula: ~1 | plotID %in% Park
(Intercept) Residual

StdDev: 3.3702 5.133291

- ▶ Значения высоты травяного покрова похожи внутри участка. Сходство наблюдений внутри одного парка слабее.



Результаты полной модели

summary(MN1_fin)

```
# Linear mixed-effects model fit by REML
# Data: graz
#      AIC      BIC    logLik
# 1709.911 1766.758 -838.9555
#
# Random effects:
# Formula: ~1 | Park
#      (Intercept)
# StdDev:      1.574143
#
# Formula: ~1 | plotID %in% Park
#      (Intercept) Residual
# StdDev:      3.3702 5.133291
#
# Fixed effects: height ~ graze_f * AspectCat + year_f + nativecov_sq + slope
#
#              Value Std.Error DF   t-value p-value
# (Intercept)  18.449084   3.092327 227   5.966084  0.0000
# graze_f1     -5.254329   2.713106  28  -1.936647  0.0629
# AspectCatS    8.571649   2.755614  28   3.110613  0.0043
# year_f2005    6.807501   1.423865 227   4.781003  0.0000
# year_f2006    4.079720   1.416165 227   2.880822  0.0043
# year_f2007   -0.587342   1.422606 227  -0.412863  0.6801
# year_f2008   -2.703761   1.409993 227  -1.917570  0.0564
# year_f2009   -2.691918   1.409963 227  -1.909211  0.0575
# year_f2010   -2.691918   1.409963 227  -1.909211  0.0575
# year_f2011    3.729040   1.418901 227   2.628118  0.0092
# nativecov_sq  -0.531161   0.378467 227  -1.403456  0.1618
# slope        -0.028216   0.052461  28  -0.537844  0.5949
# graze_f1:AspectCatS -9.444927  3.217544  28  -2.935447  0.0066
#
# Correlation:
#
#      (Intr)  grz_f1  AspcCS  y_2005  y_2006  y_2007  y_2008
# graze_f1    -0.529
# AspectCatS  -0.680  0.756
# year_f2005  -0.258  0.002 -0.025
# year_f2006  -0.283 -0.010 -0.010  0.620
# year_f2007  -0.299 -0.018  0.000  0.608  0.623
# year_f2008  -0.276 -0.009 -0.021  0.625  0.627  0.623
# year_f2009  -0.277 -0.010 -0.021  0.625  0.627  0.623  0.632
# year_f2010  -0.277 -0.010 -0.021  0.625  0.627  0.623  0.632
# year_f2011  -0.301 -0.022 -0.005  0.608  0.625  0.631  0.626
# nativecov_sq -0.202 -0.100  0.125 -0.105  0.012  0.096 -0.017
# slope       -0.423 -0.261 -0.032  0.007  0.004  0.001  0.005
# graze_f1:AspectCatS 0.500 -0.815 -0.839  0.011  0.013  0.014  0.018
```



Данные для графика предсказаний фиксированной части модели

Используем для визуализации модель, подобранную при помощи REML

Исходные данные

```
NewData_MN1_fin <- expand.grid(graза_f = levels(граз$гразе_f),  
                             AspectCat = levels(граз$AspectCat),  
                             year_f = levels(граз$year_f))  
NewData_MN1_fin$nativecov_sq <- mean(граз$nativecov_sq)  
NewData_MN1_fin$slope <- mean(граз$slope)
```

Предсказанные значения при помощи матриц

```
X <- model.matrix(~ граза_f * AspectCat + year_f + nativecov_sq + slope,  
                 data = NewData_MN1_fin)  
betas = fixef(MN1_fin)  
NewData_MN1_fin$fitted <- X %*% betas
```

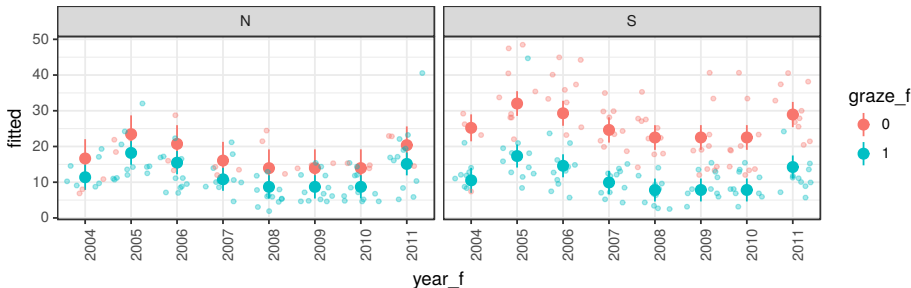
Стандартные ошибки и дов. интервалы

```
NewData_MN1_fin$se <- sqrt( diag(X %*% vcov(MN1_fin) %*% t(X)) )  
NewData_MN1_fin$lwr <- NewData_MN1_fin$fit - 1.96 * NewData_MN1_fin$se  
NewData_MN1_fin$upr <- NewData_MN1_fin$fit + 1.96 * NewData_MN1_fin$se
```

График предсказаний фиксированной части модели

На южных склонах высота травы выше там, где не пасут скот, а на северных нет. (Строго говоря, нужен еще пост хок тест, чтобы это утверждать.)

```
ggplot(data = NewData_MN1_fin, aes(x = year_f, y = fitted, colour = graze_f)) +  
  geom_pointrange(aes(ymin = lwr, ymax = upr)) +  
  facet_wrap(~ AspectCat) +  
  geom_jitter(data = graz, aes(y = height), alpha = 0.35, size = 1) +  
  theme(axis.text.x = element_text(angle = 90))
```



Вариант решения с подбором оптимальной модели
(самостоятельно)

Задание 4

- Оптимизируйте модель с предыдущего шага
- Сделайте анализ остатков
- Опишите и визуализируйте финальную модель

Решение: Подбор оптимальной модели (1)

Для подбора оптимальной модели воспользуемся тестами отношения правдоподобий. Для него нужно использовать модели, подобранные при помощи ML

```
drop1(MN1, test = "Chi")
```

```
# Single term deletions
#
# Model:
# height ~ graze_f * AspectCat + year_f + nativecov_sq + slope
#
#           Df    AIC      LRT   Pr(>Chi)
# <none>          1729.4
# year_f          7 1819.8 104.405 < 2.2e-16 ***
# nativecov_sq     1 1728.9   1.478  0.224030
# slope            1 1727.8   0.450  0.502531
# graze_f:AspectCat 1 1736.3   8.956  0.002765 **
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Решение: Подбор оптимальной модели (2)

```
MN1.1 <- update(MN1, .~.-slope)
drop1(MN1.1, test = "Chi")
```

```
# Single term deletions
```

```
#
```

```
# Model:
```

```
# height ~ graze_f + AspectCat + year_f + nativecov_sq + graze_f:AspectCat
```

```
#               Df      AIC      LRT  Pr(>Chi)
```

```
# <none>                1727.8
```

```
# year_f                7 1818.1 104.277 < 2.2e-16 ***
```

```
# nativecov_sq          1 1727.5   1.681  0.194802
```

```
# graze_f:AspectCat     1 1734.4   8.535  0.003484 **
```

```
# ---
```

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

Решение: Подбор оптимальной модели (3)

```
MN1.2 <- update(MN1.1, .~.-nativecov_sq)
drop1(MN1.2, test = "Chi")
```

```
# Single term deletions
```

```
#
```

```
# Model:
```

```
# height ~ graze_f + AspectCat + year_f + graze_f:AspectCat
```

```
#               Df      AIC      LRT  Pr(>Chi)
```

```
# <none>                1727.5
```

```
# year_f                7 1818.7 105.152 < 2.2e-16 ***
```

```
# graze_f:AspectCat    1 1733.6   8.087  0.004458 **
```

```
# ---
```

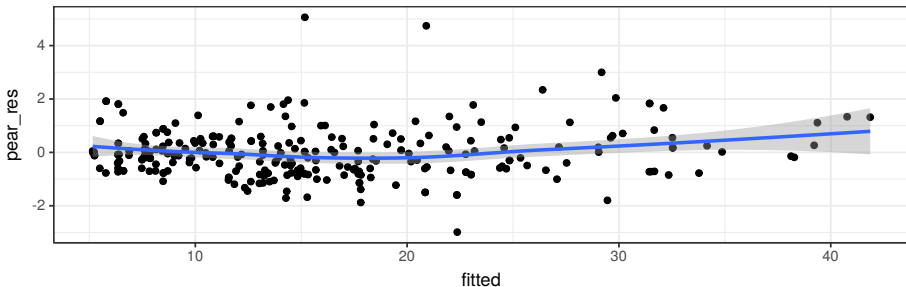
```
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Решение: Анализ остатков

```
# Данные для анализа остатков  
MN1.2_diag <- data.frame(  
  graz,  
  pear_res = residuals(MN1.2, type = "pearson"),  
  fitted = fitted(MN1.2, type = "response"))
```

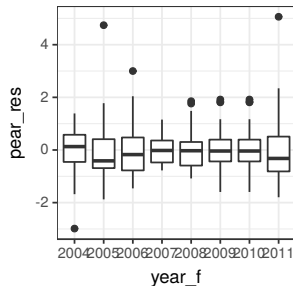
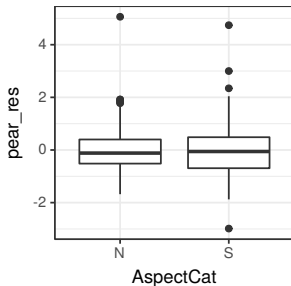
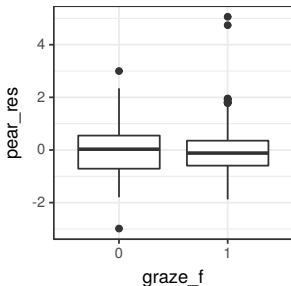

Решение: График остатков

```
gg_res <- ggplot(data = MN1.2_diag, aes(y = pear_res))  
gg_res + geom_point(aes(x = fitted)) +  
  geom_smooth(aes(x = fitted))
```



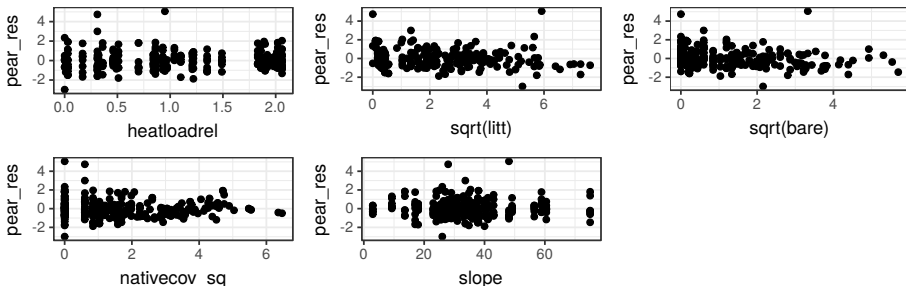
Решение: Графики остатков от переменных в модели

```
grid.arrange(gg_res + geom_boxplot(aes(x = graze_f)),  
gg_res + geom_boxplot(aes(x = AspectCat)),  
gg_res + geom_boxplot(aes(x = year_f)),  
ncol = 3)
```



Решение: Графики остатков от переменных не в модели

```
grid.arrange(  
  gg_res + geom_point(aes(x = heatloadrel)),  
  gg_res + geom_point(aes(x = sqrt(litt))),  
  gg_res + geom_point(aes(x = sqrt(bare))),  
  gg_res + geom_point(aes(x = nativecov_sq)),  
  gg_res + geom_point(aes(x = slope)),  
  ncol = 3)
```



Паттерн на графике `heatloadrel`,
`nativecov_sq`

Решение: Тестируем влияние факторов в финальной модели

Для тестов отношения правдоподобий используем финальную модель, подобранную при помощи ML

Anova(MN1.2)

```
# Analysis of Deviance Table (Type II tests)
#
# Response: height
#           Chisq Df Pr(>Chisq)
# graze_f      69.4323 1 < 2.2e-16 ***
# AspectCat     3.5606 1  0.059165 .
# year_f      131.9442 7 < 2.2e-16 ***
# graze_f:AspectCat 10.7892 1  0.001021 **
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Высота растительного покрова:

- ▶ на склонах разной экспозиции по-разному зависит от выпаса скота (достоверное взаимодействие)
- ▶ различается в разные годы
- ▶ не зависит от покрытия местных растений и крутизны склона



Решение: Описываем результаты

Для описания результатов используем модель, подобранную при помощи REML, т.к. он дает более точные оценки случайных эффектов

```
MN1.2_fin <- update(MN1.2, method = "REML")
```

Решение: Результаты

summary(MN1.2_fin)

```
# Linear mixed-effects model fit by REML
# Data: graz
#      AIC      BIC    logLik
# 1703.635 1753.485 -837.8177
#
# Random effects:
# Formula: ~1 | Park
# (Intercept)
# StdDev:    1.176709
#
# Formula: ~1 | plotID %in% Park
# (Intercept) Residual
# StdDev:    3.709642 5.090746
#
# Fixed effects: height ~ graze_f + AspectCat + year_f + graze_f:AspectCat
#
#              Value Std.Error DF   t-value p-value
# (Intercept)  16.407668  2.776747 228   5.908952  0.0000
# graze_f1     -5.678094  2.760273  29  -2.057077  0.0488
# AspectCatS    9.440473  2.875011  29   3.283630  0.0027
# year_f2005    6.626295  1.404286 228   4.718622  0.0000
# year_f2006    4.131443  1.404286 228   2.942024  0.0036
# year_f2007   -0.368337  1.404286 228  -0.262295  0.7933
# year_f2008   -2.704054  1.398243 228  -1.933894  0.0544
# year_f2009   -2.689610  1.398243 228  -1.923564  0.0557
# year_f2010   -2.689610  1.398243 228  -1.923564  0.0557
# year_f2011    3.987275  1.398243 228   2.851633  0.0047
# graze_f1:AspectCatS -9.642549  3.372692  29  -2.859007  0.0078
# Correlation:
#
# (Intr) grz_f1 AspcCS y_2005 y_2006 y_2007 y_2008
# graze_f1      -0.807
# AspectCatS    -0.787  0.795
# year_f2005    -0.306 -0.011 -0.015
# year_f2006    -0.306 -0.011 -0.015  0.625
# year_f2007    -0.306 -0.011 -0.015  0.625  0.625
# year_f2008    -0.305 -0.012 -0.021  0.627  0.627  0.627
# year_f2009    -0.305 -0.012 -0.021  0.627  0.627  0.627  0.632
# year_f2010    -0.305 -0.012 -0.021  0.627  0.627  0.627  0.632
# year_f2011    -0.305 -0.012 -0.021  0.627  0.627  0.627  0.632
# graze_f1:AspectCatS 0.654 -0.825 -0.849  0.016  0.016  0.016  0.021
#
# y_2009 y_2010 y_2011
# graze_f1
# AspectCatS
# year_f2005
```



Решение: Внутрикласовая корреляция

Для расчета нужна модель, подобранная при помощи REML

```
MN1.2_fin
```

Для наблюдений на одном и том же участке $\sigma_{plotID}^2 / (\sigma_{plotID}^2 + \sigma_{Park}^2 + \sigma^2)$

```
3.709642^2 / (1.176709^2 + 3.709642^2 + 5.090746^2)
```

```
# [1] 0.3351399
```

Для наблюдений в одном и том же парке $\sigma_{Park}^2 / (\sigma_{plotID}^2 + \sigma_{Park}^2 + \sigma^2)$

```
1.176709^2 / (1.176709^2 + 3.709642^2 + 5.090746^2)
```

```
# [1] 0.03372099
```

В результатах

Random effects:

```
Formula: ~1 | Park  
(Intercept)
```

```
StdDev: 1.176709
```

```
Formula: ~1 | plotID %in% Park  
(Intercept) Residual
```

```
StdDev: 3.709642 5.090746
```

Решение: Данные для графика предсказаний фиксированной части модели

Используем для визуализации модель, подобранную при помощи REML

Исходные данные

```
NewData_MN1.2_fin <- expand.grid(graза_f = levels(graза$газа_f),  
                                AspectCat = levels(graза$AspectCat),  
                                year_f = levels(graза$year_f))  
NewData_MN1.2_fin$nativecov_sq <- mean(graза$nativecov_sq)  
NewData_MN1.2_fin$slope <- mean(graза$slope)
```

Предсказанные значения при помощи матриц

```
X <- model.matrix(~ граза_f * AspectCat + year_f, data = NewData_MN1.2_fin)  
betas = fixef(MN1.2_fin)  
NewData_MN1.2_fin$fitted <- X %*% betas
```

Стандартные ошибки и дов. интервалы

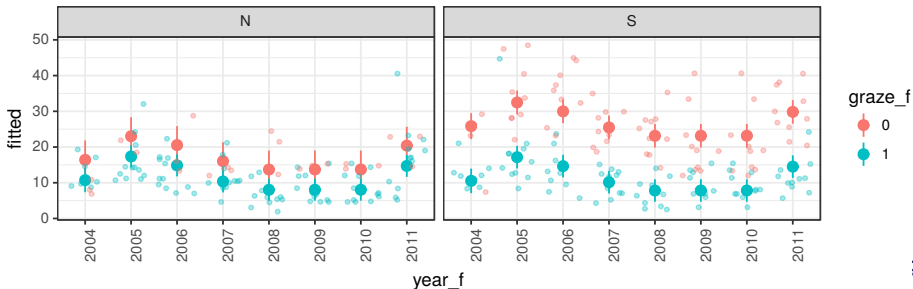
```
NewData_MN1.2_fin$se <- sqrt( diag(X %*% vcov(MN1.2_fin) %*% t(X)) )  
NewData_MN1.2_fin$lwr <- NewData_MN1.2_fin$fit - 1.96 * NewData_MN1.2_fin$se  
NewData_MN1.2_fin$upr <- NewData_MN1.2_fin$fit + 1.96 * NewData_MN1.2_fin$se
```


Решение: График предсказаний фиксированной части модели

На южных склонах высота травы выше там, где не пасут скот, а на северных нет. (Строго говоря, нужен еще пост хок тест, чтобы это утверждать)

Не удивляйтесь тому, что график похож на предыдущий, т.к. те факторы, которые мы удалили из модели и так не влияли.

```
ggplot(data = NewData_MN1.2_fin, aes(x = year_f, y = fitted, colour = graze_f)) +  
  geom_pointrange(aes(ymin = lwr, ymax = upr)) +  
  facet_wrap(~ AspectCat) +  
  geom_jitter(data = graz, aes(y = height), alpha = 0.35, size = 1) +  
  theme(axis.text.x = element_text(angle = 90))
```



Take-home messages

- ▶ Случайные факторы в смешанных моделях могут быть вложены друг в друга
- ▶ Есть два способа подбора коэффициентов в смешанных моделях: ML и REML. Для разных этапов анализа важно, каким именно способом подобрана модель.

- ▶ Crawley, M.J. (2007). The R Book (Wiley).
- ▶ Zuur, A. F., Hilbe, J., & Ieno, E. N. (2013). A Beginner's Guide to GLM and GLMM with R: A Frequentist and Bayesian Perspective for Ecologists. Highland Statistics.
- ▶ Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., and Smith, G.M. (2009). Mixed Effects Models and Extensions in Ecology With R (Springer)
- ▶ Pinheiro, J., Bates, D. (2000). Mixed-Effects Models in S and S-PLUS. Springer