MODEL CHECKS FOR FOUR FEEDING STUDIES

CLASS 9 · DIAGNOSTIC WORKSHOP

Gavin Simpson
Aarhus University

Wednesday, 4 November 2026

DIET COMPARISONS AND DIAGNOSTIC QUESTIONS

Case	Data and starting model	Question
1	Event counts with Poisson and negative binomial fits	Can either model support a proposed reduction in counts?
2	Animal growth; ordinary linear model	Do animals share pen differences?
3	Growth and starting weight	Is the fitted weight relationship suitable?
4	Positive/negative welfare screen	Can we estimate a reliable diet comparison?

Choose a case worksheet. Try the formula and read the output before opening the solutions.

EVIDENCE AND POSSIBLE CAUSES BEFORE THE BREAK

- 10 minutes: identify who received diet and what was measured.
- 15 minutes: inspect model output and plots. Use the morning's fits for pen counts.
- 15 minutes: discuss two possible causes of any concern.
- 10 minutes: justify a model choice or identify what further information you need.
- 10 minutes: exchange an explanation with another group.

Keep **one concern and one reassuring feature**, and say what each can tell you.

CHECKS FOR SHARED PENS, STARTING WEIGHT AND WELFARE SCREENS

In the shared-pen case, six animals live in each pen. A bell-shaped residual plot can still hide shared pen differences.

In the growth case, researchers measured starting weight before assigning diets. Check unexplained gain against that weight, even if another plot looks reassuring.

In the welfare-screen case, compare the observed counts by diet with the fitted probabilities and the diet coefficient.

For your case, which extra output or plot addresses the design concern?

EXPLAIN YOUR EVIDENCE TO ANOTHER GROUP

Tell the paired group about your case's measured outcome and diet assignment first. Then explain:

"Because [output or pattern], we suspect [problem]. [Another cause] could also explain it. Our design suggests [action]. We will check [evidence] again."

The listener names one claim that needs more evidence.

Break 14:55–15:10.

SHARED-PEN CASE: READ AND PREPARE THE DATA

Researchers assigned diets to 16 pens in four rooms. Each room had two pens per diet. They measured gain in g/day for all six animals in each pen, giving 96 measurements.

```
library("dplyr")
d <- readr::read_csv("data/case-2.csv")
d <- d |>
  mutate(
    diet = factor(diet, levels = c("Control", "High fibre")),
    room = factor(room, levels = c("R1", "R2", "R3", "R4")),
    pen = factor(pen)
  )
d |>
  select(room, pen, diet, gain_g_day) |>
  head(2)
```

```
# A tibble: 2 × 4
  room  pen    diet    gain_g_day
  <fct> <fct>  <fct>        <dbl>
1 R1    R1-P01 Control       883.
2 R1    R1-P01 Control       819.
```

MEAN GAIN WITH AND WITHOUT A SHARED PEN EFFECT

Both models estimate a High fibre minus Control difference in mean gain while allowing the four rooms to differ. Read `(1 | pen)` as "a random intercept for each pen". As in Class 8, it shifts expected gain up or down by the same amount for every animal in a pen. The model estimates the spread of these pen differences.

```
library("lme4")

original <- lm(gain_g_day ~ diet + room, data = d)

revised <- lmer(
  gain_g_day ~ diet + room + (1 | pen),
  data = d,
  REML = TRUE
)
```

REML estimates pen and animal variation while accounting for the fitted diet and room means.

DID THE PEN MODEL FIT NUMERICALLY?

```
library("performance")

check_convergence(revised)
```

```
[1] TRUE
attr(,"gradient")
[1] 1.940595e-08
```

```
check_singularity(revised)
```

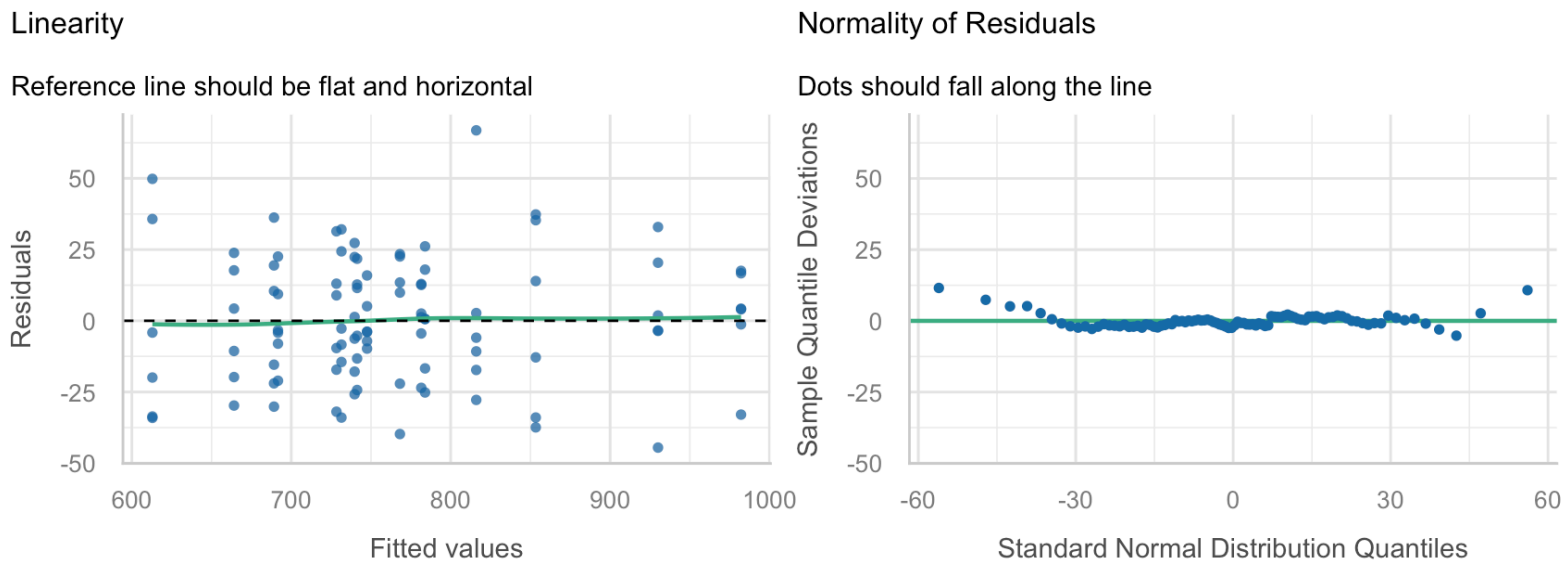
```
[1] FALSE
```

The fitting check returns `TRUE`. The singularity check returns `FALSE`, which here means the fitted spread of pen differences is above zero.

RESIDUAL CHECKS FOR THE SHARED-PEN MODEL

QQ compares residual shape with a bell shape. The other panel looks for unexplained patterns across fitted gain.

```
plot(check_model(revised, check = c("qq", "linearity")))
```



Your script also plots residuals by pen. Why is that extra view useful when six animals share conditions?

WHICH ANIMALS DOES THE GAIN COMPARISON REPRESENT?

For each recorded animal, `avg_comparisons()` predicts gain under both diets in its room, subtracts Control from High fibre, then averages the differences in g/day. The rooms get equal weight because each has the same number of records. A balanced grid would give the rooms equal weight by construction.

For the shared-pen model, `re.form = NA` sets the fitted pen differences to zero. The comparison describes the diets across the four rooms at the average pen effect.

DIET GAIN DIFFERENCE FROM THE ORDINARY LINEAR MODEL

```
library("marginaleffects")

avg_comparisons(original, variables = "diet")
```

```
 Estimate Std. Error    z Pr(>|z|)    S 2.5 % 97.5 %
     61.5       14.9 4.14   <0.001 14.8  32.4   90.6

Term: diet
Type: response
Comparison: High fibre - Control
```

The estimate is **61.5 g/day**, with an approximate 95% interval of **32.4 to 90.6 g/day**. This fit treats the 96 animals' remaining gain differences as independent.

THE SHARED-PEN MODEL GIVES A WIDER INTERVAL

```
avg_comparisons(revised, variables = "diet", re.form = NA)
```

```
 Estimate Std. Error    z Pr(>|z|)   S 2.5 % 97.5 %
     61.5       40.7 1.51     0.13 2.9 -18.2    141

Term: diet
Type: response
Comparison: High fibre - Control
```

The estimate is still **61.5 g/day**. The same number of pens and animals contribute to each diet within each room. The shared-pen model accounts for the animals sharing conditions and for the 16 pen-level diet assignments.

Its approximate 95% interval is **−18.2 to 141.2 g/day**. This interval includes zero and allows for a decrease as well as an increase in gain.

RECHECK THE ANALYSIS AFTER THE BREAK

Return to your current case or choose another worksheet.

- **Pen counts.** Compare both intervals with 0.80, then 0.70. Keep the fits and room weights fixed. Explain which reduction you could support.
- **Growth and welfare screens.** Explain the proposed follow-up. Use the output to say what changed and what you can report.

CONFIRM THE PROJECT QUESTION AND USABLE DATA

For the group project, name the primary outcome, main comparison and unit that received treatment. Confirm that you can open the data and understand the design.

If essential data or design information is missing, use the supplied project case.

REPORT AND INDIVIDUAL EXIT

At 15:50, each group has one minute to report on a case. Follow the five report steps: study and starting model, evidence and possible causes, action and reason, recheck, and biological conclusion and limitation.

For your individual exit, imagine a welfare-screen model returns “converged” and almost no unexplained variation. All Control animals screened positive and all High fibre animals negative. What else would you inspect before trusting a diet estimate?

Next class: several measurements from the same animal. Bring your report, trajectory sketch and one question.