FOLLOW UP THE EVIDENCE

CLASS 9: REPORT AND PREPARE FOR REPEATED MEASUREMENTS

Gavin Simpson

Aarhus University

Wednesday, 4 November 2026

AFTER CLASS: 60–75 MINUTES

- 25 minutes: the recording or these slides.
- 25–35 minutes: targeted Touchon reading.
- 10–15 minutes: finish your report and draw animal trajectories.

Bring your report, sketch and one question to Class 10.

FIVE STEPS IN THE DIAGNOSTIC REPORT

1. **Study and starting model.** Say how researchers assigned the diets and what they measured.
2. **Evidence and possible causes.** Identify a concern, a reassuring feature and two possible explanations.
3. **Action and reason.** Explain the change you chose.
4. **Recheck.** Compare the new output with the original.
5. **Biological conclusion and limitation.** Say what the evidence supports and what remains uncertain.

PEN COUNTS IN FOUR ROOMS

Researchers assigned diets to pens, then counted treatment events once per pen over the same follow-up time. In the model, `diet` compares the diets and `room` allows the four rooms to differ.

```
library("dplyr")
d <- readr::read_csv("data/case-1.csv")
d <- d |>
  mutate(
    diet = factor(diet, levels = c("Control", "High fibre")),
    room = factor(room, levels = c("R1", "R2", "R3", "R4"))
  )
head(d, 2)
```

```
# A tibble: 2 × 4
  room  pen    diet    events
  <fct> <chr>  <fct>    <dbl>
1 R1    R1-P01 Control      1
2 R1    R1-P02 Control      1
```

TWO MODELS FOR THE SAME PEN COUNTS

Both models estimate the expected count for each diet and room. The Poisson model sets the count variance equal to the expected count. The negative-binomial model lets counts vary more around that expectation.

```
poisson_fit <- glm(
  events ~ diet + room,
  family = poisson(),
  data = d
)
library("mgcv")
```

```
Loading required package: nlme

Attaching package: 'nlme'

The following object is masked from 'package:dplyr':

    collapse

This is mgcv 1.9-4. For overview type '?mgcv'.
```

```
nb_fit <- gam(
  events ~ diet + room,
  family = nb(),
  data = d,
  method = "ML"
)
```

THE POISSON COUNT RATIO

`newdata = "balanced"` gives each of the four rooms equal weight. The observed rows give the same weights because each room has the same number of pens.

```
library("marginaleffects")
avg_comparisons(
  poisson_fit,
  variables = "diet",
  newdata = "balanced",
  comparison = "lnratioavg",
  transform = "exp"
)
```

```
 Estimate Pr(>|z|)   S 2.5 % 97.5 %
     0.65  0.00153 9.3 0.497  0.848

Term: diet
Type: response
Comparison: ln(mean(High fibre) / mean(Control))
```

R averages the room predictions for each diet, then takes the log of the High fibre mean divided by the Control mean. `transform = "exp"` returns the estimate and interval to the ratio scale.

THE COUNT RATIO WITH EXTRA VARIATION

```
avg_comparisons(
  nb_fit,
  variables = "diet",
  newdata = "balanced",
  comparison = "lnratioavg",
  transform = "exp"
)
```

```
 Estimate Pr(>|z|)   S 2.5 % 97.5 %
    0.673    0.133 2.9 0.401   1.13

Term: diet
Type: response
Comparison: ln(mean(High fibre) / mean(Control))
```

Both estimates suggest fewer events under High fibre. The Poisson interval lies below 1, while the wider interval here includes 1. Under the model that allows extra variation, the evidence also allows equal expected counts.

DOES THE NEGATIVE-BINOMIAL MODEL ALLOW ENOUGH VARIATION?

```
library("performance")
set.seed(9401)
check_overdispersion(nb_fit, iterations = 500)
```

```
# Overdispersion test (using simulated residuals)

 dispersion ratio = 0.868
          p-value =  0.78
```

`iterations = 500` asks for 500 simulated datasets. `set.seed()` makes that check repeatable. The ratio is close to 1, which gives little evidence of extra variation beyond what this model allows. Reopen the pen-count case to compare the residual plots.

WELFARE SCREENS SHOW COMPLETE SEPARATION

The welfare-screen case has animals housed separately, one screen per animal. Read the observed table before trusting the logistic output.

```
screen <- readr::read_csv("data/case-4.csv")
screen |>
  summarise(
    positive = sum(screened_positive == 1),
    negative = sum(screened_positive == 0),
    .by = diet
  )
```

```
# A tibble: 2 × 3
  diet       positive negative
  <chr>         <int>    <int>
1 Control          12        0
2 High fibre        0       12
```

All 12 Control animals screened positive. All 12 High fibre animals screened negative. Diet perfectly predicts the observed screens, a pattern called **complete separation**. An ordinary logistic fit can report convergence even though it has no finite room-adjusted diet estimate. Report the observed rates and their intervals.

REPEATED GAIN MEASUREMENTS IN TWO PENS

Researchers assigned one diet to each pen. They recorded each animal's gain in grams per day (`gain_g_day`) at weeks 0–3. The data contain six animals and four measurements per animal.

```
bridge <- readr::read_csv("data/repeated-measures-bridge.csv")
bridge |> distinct(pen, diet)
```

```
# A tibble: 2 × 2
  pen    diet
  <chr>  <chr>
1 R1-P01 Control
2 R1-P02 High fibre
```

Each row in the original data is one animal at one week. The two pens received the diets once each.

GAINS FOR THE THREE ANIMALS IN YOUR SKETCH

Open the repeated-measurements worksheet. `filter()` keeps the three named animals and `pivot_wider()` gives each animal its own column.

```
bridge |>
  filter(animal %in% c("R1-P01-A1", "R1-P01-A2", "R1-P02-A1")) |>
  select(time_week, animal, gain_g_day) |>
  tidyr::pivot_wider(
    names_from = animal,
    values_from = gain_g_day
  )
```

```
# A tibble: 4 × 4
  time_week `R1-P01-A1` `R1-P01-A2` `R1-P02-A1`
      <dbl>       <dbl>       <dbl>       <dbl>
1         0        722.        679.        720.
2         1        715.        703.        752.
3         2        727.        711.        766.
4         3        764.        727.        773.
```

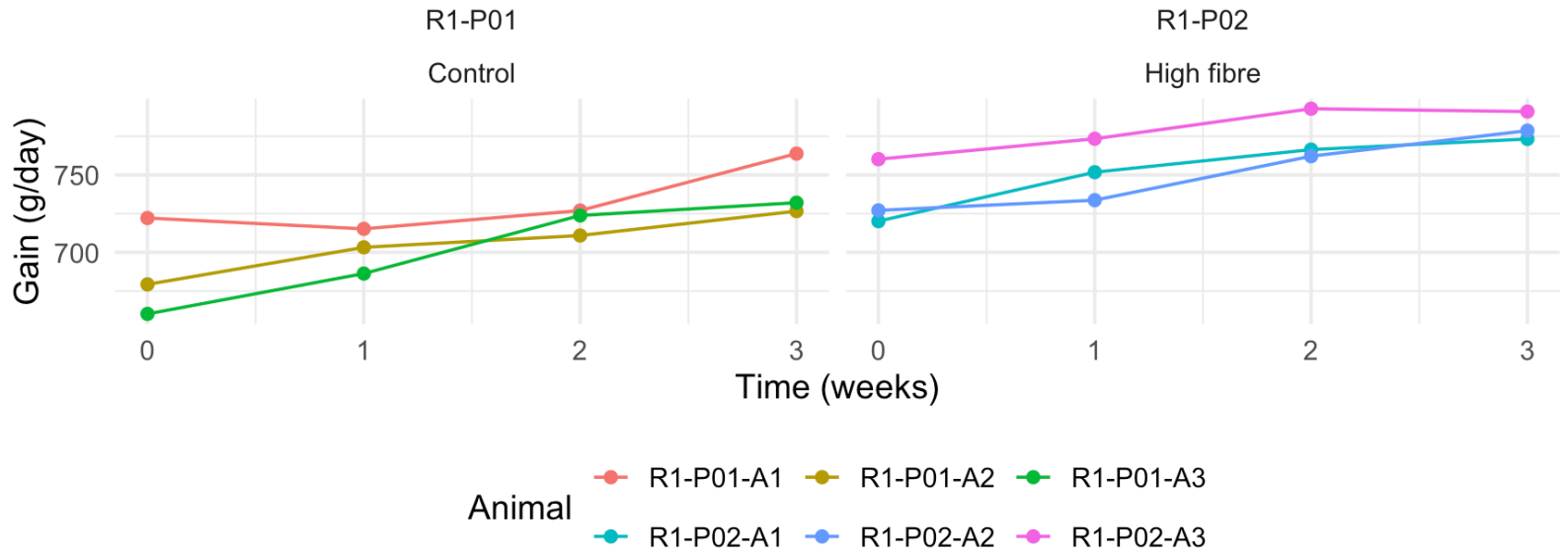
Sketch these three trajectories at weeks 0–3. Label the axes with week and gain (g/day), then mark each animal's pen and diet.

ANIMAL TRAJECTORIES BY PEN AND DIET

Each point is an observed gain. Each line connects measurements from the same animal. The panel headings give the pen and its diet. With one pen per diet, differences between pens could reflect diet or other conditions shared within each pen.

```
library("ggplot2")
ggplot(
  bridge,
  aes(time_week, gain_g_day, group = animal, colour = animal)
) +
  geom_line() +
  geom_point() +
  facet_wrap(vars(pen, diet), nrow = 1) +
  scale_x_continuous(breaks = 0:3) +
  labs(x = "Time (weeks)", y = "Gain (g/day)", colour = "Animal") +
  guides(colour = guide_legend(nrow = 2, byrow = TRUE)) +
  theme_minimal(base_size = 14) +
  theme(legend.position = "bottom")
```

ANIMAL TRAJECTORIES BY PEN AND DIET



FINISH YOUR REPORT AND READING

Follow the five steps in the diagnostic report handout and bring your report and sketch to Class 10. Which points come from the same animal? Which animals share a pen? Write one question about whether the diet difference changes over time, and suggest why nearby measurements from one animal might be alike.

In Touchon's second edition, revisit model checking in Chapter 7 or 8. For the starting-weight case, revisit mean relationships in Chapter 6. Then read the introduction to repeated measurements in Chapter 9.