POWER AND MODEL ASSUMPTIONS

AFTER WEEK 6: CONSOLIDATION AND THE WEEK 7 BRIDGE

Gavin Simpson
Aarhus University

Wednesday, 21 October 2026

AFTER-CLASS ROUTE

Within the **150-minute autumn-break task**, use 30 minutes for revision:

Time	Focus
10 minutes	Reconstruct generation and fitting
10 minutes	Compare supplied baseline and sensitivity results
10 minutes	Annotate your model map and record a question

Use relevant slides selectively. Detailed separation, overdispersion and precision are optional depth; no new fitting is required.

WHAT THE CURVE ASSUMES

Each point depends on a complete proposed experiment:

- response family, link and baseline
- treatment effect and remaining variation
- units, allocation and fixed room effects
- fitted model and rejection rule

The error bars on our graph do not express uncertainty about all these assumptions.

QUESTION: THE SAME ODDS RATIO

Two binary scenarios use an odds ratio of 0.60.

The control probability is 0.30 in one and 0.15 in the other.

Are their absolute probability differences equal? Should we expect their power curves to be identical?

ANSWER: BASELINE PROBABILITY MATTERS

At control probability **0.30**, the high-fibre probability is about **0.205**: a difference of about **−0.095**.

At control probability **0.15**, it is about **0.096**: a difference of about **−0.054**.

The odds ratio is unchanged, but response-scale differences and information differ.

EFFECT RELEVANCE AND UNCERTAINTY

Choose an effect worth detecting using scientific evidence and consequences.

Do not substitute the largest plausible effect just to obtain a smaller sample size. Investigate an effect range when evidence is uncertain.

Our two scenarios illustrate sensitivity; they do not exhaust uncertainty about the real experiment.

A QUALIFIED RECOMMENDATION

A useful recommendation states the assumptions, design, analysis, effect, criterion and uncertainty.

It can conclude that no investigated design meets the target, or that better information about variation is needed first.

Power is a design property under assumptions, not a posterior probability that a hypothesis is true.

QUESTION: MORE SIMULATION

One group uses 200 repetitions; another uses 2,000.

Both use exactly the same experiment and parameters.

Which uncertainty should decrease? Which uncertainty remains? Does the proposed experiment gain independent replication?

ANSWER: MONTE CARLO PRECISION

For an estimated detection probability near 0.80, the approximate Monte Carlo standard error is:

- **0.028** with 200 repetitions
- **0.009** with 2,000 repetitions

More repetitions sharpen the numerical estimate. They do not resolve uncertainty about effect size, response variation or feasibility.

UNRESOLVED FITS REMAIN VISIBLE

Our procedure records warnings, failed comparisons and suspicious binary fits. They are not counted as detections.

A graph of detections divided by **all attempts** describes the operational procedure. Report the unresolved fraction alongside it.

Deleting unresolved fits answers a different question and can distort the planning conclusion.

QUESTION: EXTRA COUNT VARIATION

The Poisson generator has variance μ.

A negative-binomial generator with size 1.5 has variance $\mu + \mu^2/1.5$. Suppose we keep the same means but still fit a Poisson model.

If rejection becomes more frequent, has the design improved? What should we check first?

ANSWER: INSPECT NULL REJECTION

First compare rejection when the true diet effect is **zero**.

If the nominal 5% procedure rejects much more often, a higher alternative rejection rate is not a clean gain in power at the intended error rate.

An analysis accommodating the extra variation must be checked under both null and alternative before making a recommendation.

COUNT-MODEL REFERENCE AT 64 PENS

Generator	True diet effect	Rejection	95% MC interval
Poisson	Ratio 0.65	85.7%	84.0%–87.1%
Poisson	Zero	4.7%	3.8%–5.7%
Overdispersed	Ratio 0.65	72.2%	70.2%–74.1%
Overdispersed	Zero	29.7%	27.7%–31.7%

Both generators use the **Poisson analysis**. Each percentage uses 2,000 attempts.

NULL CHECKS AND MODEL ASSUMPTIONS

A null simulation is a diagnostic of the entire declared procedure.

A departure from 5% may reflect simulation noise, small-sample approximations, screening rules or a mismatched model.

Check the interval and unresolved fraction. A null check under one scenario does not validate every alternative scenario.

BEYOND ONE FIXED ALLOCATION

In class, each sample size uses one fixed balanced allocation.

Rerandomising within rooms for every repetition instead averages over the assignment mechanism.

State which target is being estimated. Additional rooms would require new assumptions about room effects, not just a longer design table.

QUESTION: PRECISION VERSUS DETECTION

Consider these illustrative intervals for a gain difference:

- **−10 to 10 g/day**
- **5 to 85 g/day**

Which excludes zero? Which has total width at most 60 g/day?

ANSWER: DIFFERENT SUCCESS EVENTS

The first interval includes zero but is only **20 g/day wide**.

The second excludes zero but is **80 g/day wide**.

The probability of useful precision is a separate planning outcome. The 60 g/day threshold needs scientific justification in a real study.

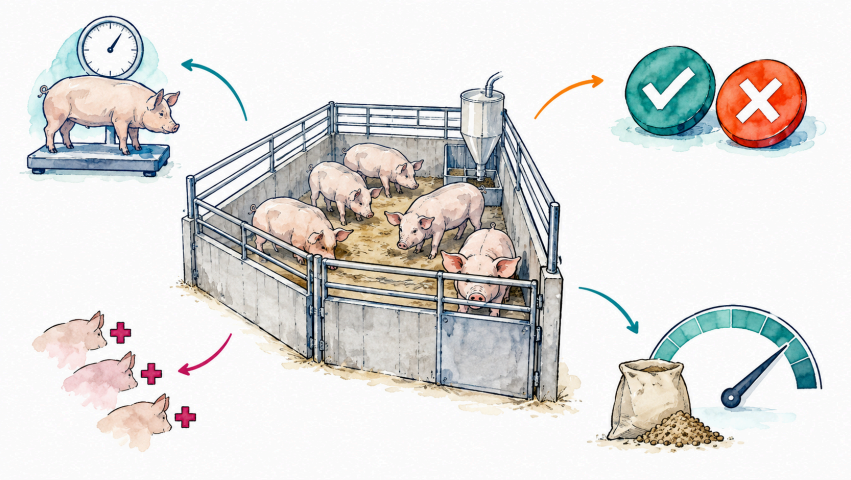
THE WEEK 7 MODEL MAP

```
response ~ diet + room
```


```
response ~ diet + room
```


Name the response, experimental unit, predictors, reference levels, family and link.

Explain how the formula represents the planned comparison.



REFERENCE LEVELS AND PREDICTIONS

The data generating process and fitted model both use **Control** and **Room 1** as reference levels.

For a fitted mean at a new pen, supply diet and room. The model does not use pen ID as a predictor.

Choose the family and link as well as the formula: the formula alone does not define a GLM.

BRING TO WEEK 7

Retain:

- your revised recommendation and its weakest assumption
- the four overdispersion comparison numbers
- a one-page model map and one question

Week 7: **21 October**. The next step is to explain and construct the model formula, building on models you have already used.

SOURCES

- R Core Team: GLMs, model comparisons, negative-binomial distribution.
- Colegrave and Ruxton (2020): *Power Analysis: An Introduction for the Life Sciences*. Instructor background.

All reference results are reproducible from the supplied teaching scenarios.