

Elab 18.7

[S2008/S2008-LOG16]

The concept of residuals does not cleanly apply to yes/no models because the model value is a probability (of a yes outcome), whereas the actual observation is the outcome itself. It would be silly to try to compute a difference between “yes” and a probability like 0.8. After all, what could it mean to calculate $(\text{yes} - 0.8)^2$?

In fitting ordinary linear models, the criterion used to select the best coefficients for any given model design is “least squares,” minimizing the sum of square residuals. The corresponding criterion in fitting yes/no models (and many other types of models) is “maximum likelihood.”

The word “likelihood” has a very specific and technical meaning in statistics, it’s not just a synonym for “chance” or “probability.” A likelihood is the probability of the outcome *according to a specific model*.

To illustrate, here is an example of some yes-no observations and the model values of two different models.

Case	Model A		Model B		Observed Outcome
	p(Yes)	p(No)	p(Yes)	p(No)	
1	0.7	0.3	0.4	0.6	Yes
2	0.6	0.4	0.8	0.2	No
3	0.1	0.9	0.3	0.7	No
4	0.5	0.5	0.9	0.1	Yes

Likelihood always refers to a given model, so there are two likelihoods here: one for Model A and another for Model B. The likelihood for each case under Model A is the probability of the observed outcome according to the model. For example, the likelihood under Model A for case 1 is 0.7, because that is the model value of the observed outcome “Yes” for that case. The likelihood of case 2 under Model A is 0.4 — that is the probability of “No” for case 2 under model A.

- What is the likelihood under Model A for case 3? 0.1 0.3 0.5 0.7 0.9 [Elab 18.7-1]
- What is the likelihood under Model B for case 3? 0.1 0.3 0.5 0.7 0.9 [Elab 18.7-2]
- What is the likelihood under Model A for case 4? 0.1 0.3 0.5 0.7 0.9 [Elab 18.7-3]
- What is the likelihood under Model B for case 4? 0.1 0.3 0.5 0.7 0.9 [Elab 18.7-4]

The likelihood for the whole set of observations combines the likelihoods of the individual cases: multiply them all together. This is justified if the cases are independent of one another, as is usually assumed and sensible if the cases are the result of random sampling or random assignment to an experimental treatment.

- What is the likelihood under Model A for the whole set of cases?

A $0.3 \times 0.4 \times 0.9 \times 0.5$

B $0.7 \times 0.6 \times 0.9 \times 0.5$

C $0.3 \times 0.4 \times 0.1 \times 0.5$

D $0.7 \times 0.4 \times 0.9 \times 0.5$

E $0.7 \times 0.4 \times 0.1 \times 0.5$

Elab 18.7-5

- What is the likelihood under Model B for the whole set of cases?

A $0.4 \times 0.8 \times 0.3 \times 0.9$

B $0.4 \times 0.2 \times 0.3 \times 0.9$

C $0.6 \times 0.2 \times 0.3 \times 0.9$

D $0.4 \times 0.2 \times 0.7 \times 0.9$

E $0.4 \times 0.2 \times 0.3 \times 0.1$

Elab 18.7-6