Model selection and goodness-of-fit testing: the basics

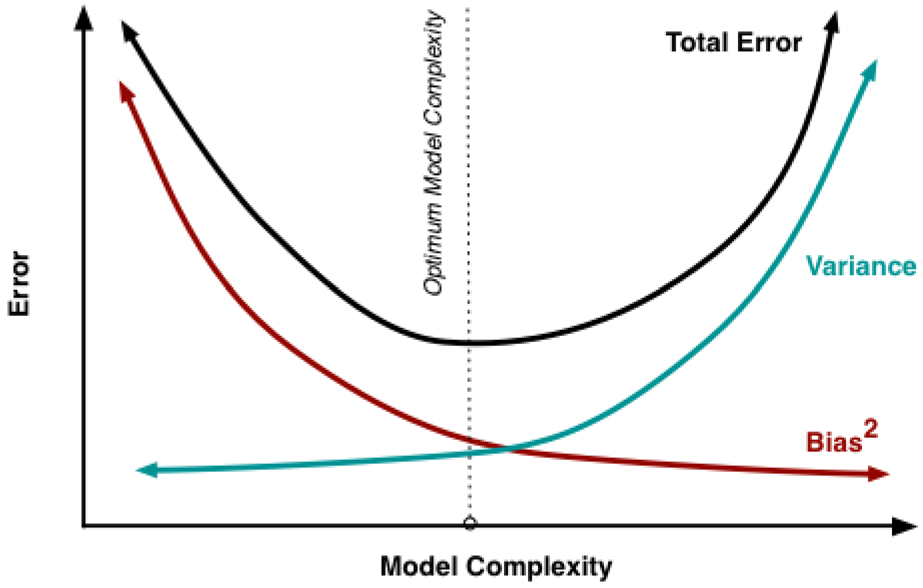
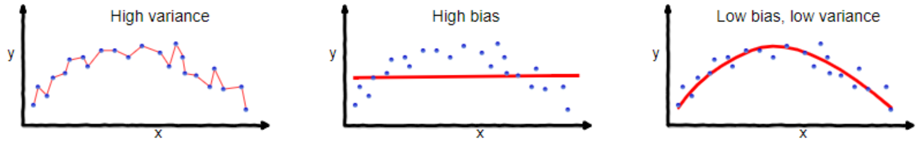


Table 1. Models examining effects of various covariates on detection probabilities of indicated breeding pairs of waterfowl in North Dakota, USA (from Pagano and Arnold 2009). I added single parameters assuming an additive effect to the base model, which included $K = 16$ parameters (8 species $\times$ 2 observers). Three of these covariates were considered biologically feasible (total ducks, vegetative cover, and cover type), 6 were not (random 5, 4, 8, and 1; not Sunday, Monday, or Wednesday; and last duck seen a mallard), and I excluded 6 additional nonsense or random variables (ΔAIC = 0.64–2.00) from presentation. I evaluated all models compared to the base model using Akaike's Information Criterion (AIC), ΔAIC, and changes in model deviance (Dev).

Model	AIC	ΔAIC	K	Dev
Total ducks	4,426.71	−16.62	17	4,392.71
Random 5	4,439.95	−3.38	17	4,405.95
Random 4	4,442.20	−1.13	17	4,408.19
Vegetative cover	4,442.65	−0.68	17	4,408.65
Random 8	4,442.81	−0.52	17	4,408.80
Random 1	4,442.90	−0.43	17	4,408.90
Base model	4,443.34	0	16	4,411.33
Not Sunday, Monday, or Wednesday	4,445.09	1.75	17	4,411.08
Cover type	4,445.25	1.92	17	4,411.25
Last duck seen a mallard	4,445.33	1.99	17	4,411.32



overfitting | underfitting | Good balance

```
> dredge(no)
```

PPPs; BCIs; CIs; p-values; DIC; AIC; BIC; AICc; LOOIC; WAIC; Bayesian p-values; deviance

Announcements

1. Next week we'll start in-class work sessions

2. Presentations and final reports (in a format of your preference) are coming up!
   - Presentations Nov 26th, Dec 3rd, and Dec 5th (< 10m)
   - Reports due Dec 5th

3. Final report formatting is extremely flexible*
   - Code and data
   - Conceptual diagram
   - Methods section (w/results?)
   - Dissertation chapter
   - Manuscript draft

***You don't have to use SEM. If you don't, just explain why in ~1 page**

Two <u>big</u> goals today (on Thursday we'll play with SEMs)

1. **Talk through model selection (using AICc as an example)**
   a) Comparing the data generating model to a null model
   b) How do meaningless covariates affect things?
   c) What happens when things get **really** confusing*?

2. **Talk through goodness-of-fit (using Bayesian p-values) using a fun but challenging example**

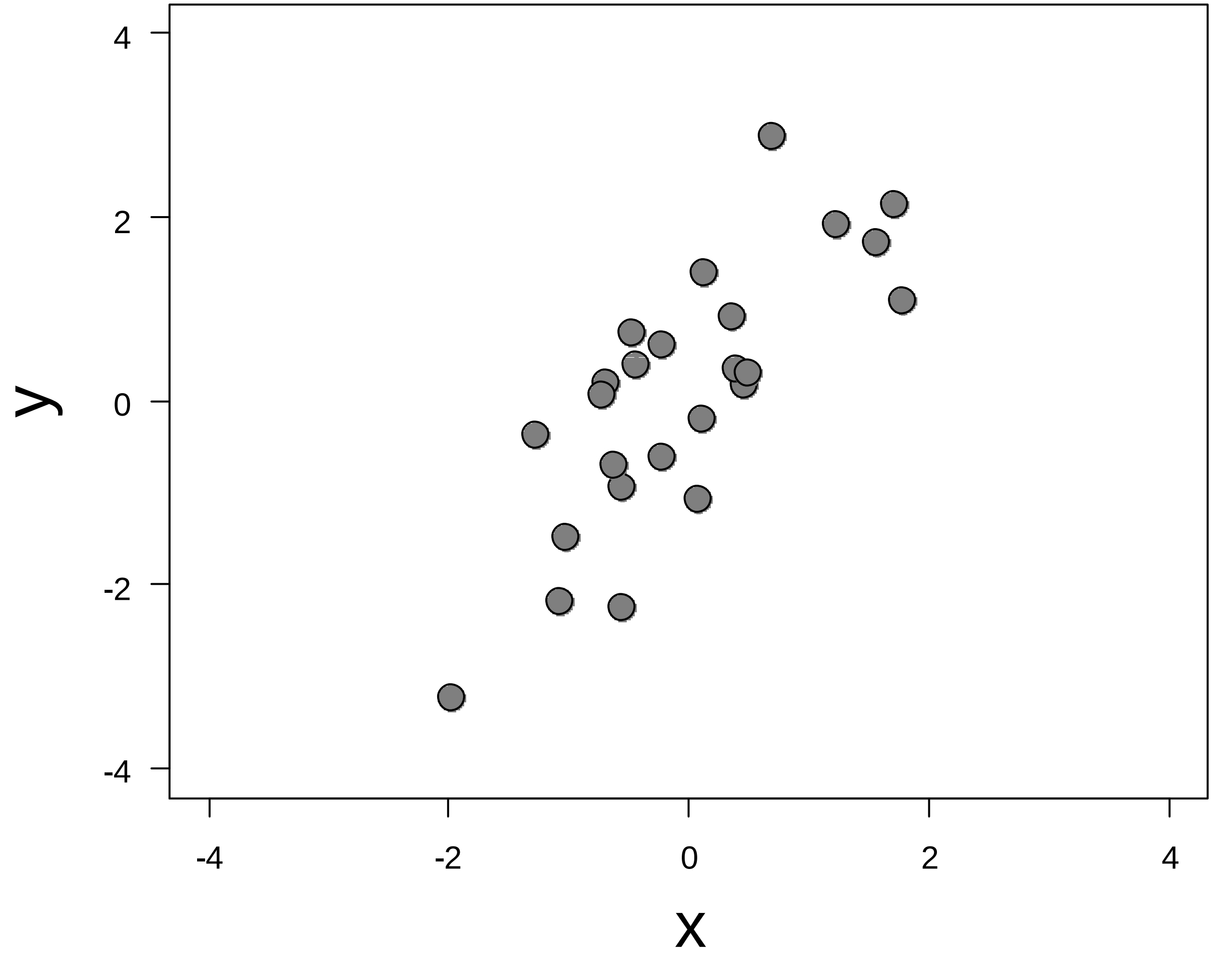
***Things will always be confusing! All models are wrong!**

How do we ‘score’ or rank models?

- Differences (δAICc) in AICc between models of < 2 indicate near equivalence
- Differences of < 7 among models indicate that some support may exist for ‘inferior’ models (i.e., just because δAICc is > 2 doesn’t mean that there is no support).

Burnham and Anderson (2002)

Simulating data (3 parameters)



Data simulation

$\boldsymbol{x} \sim \text{Normal}(0,1)$

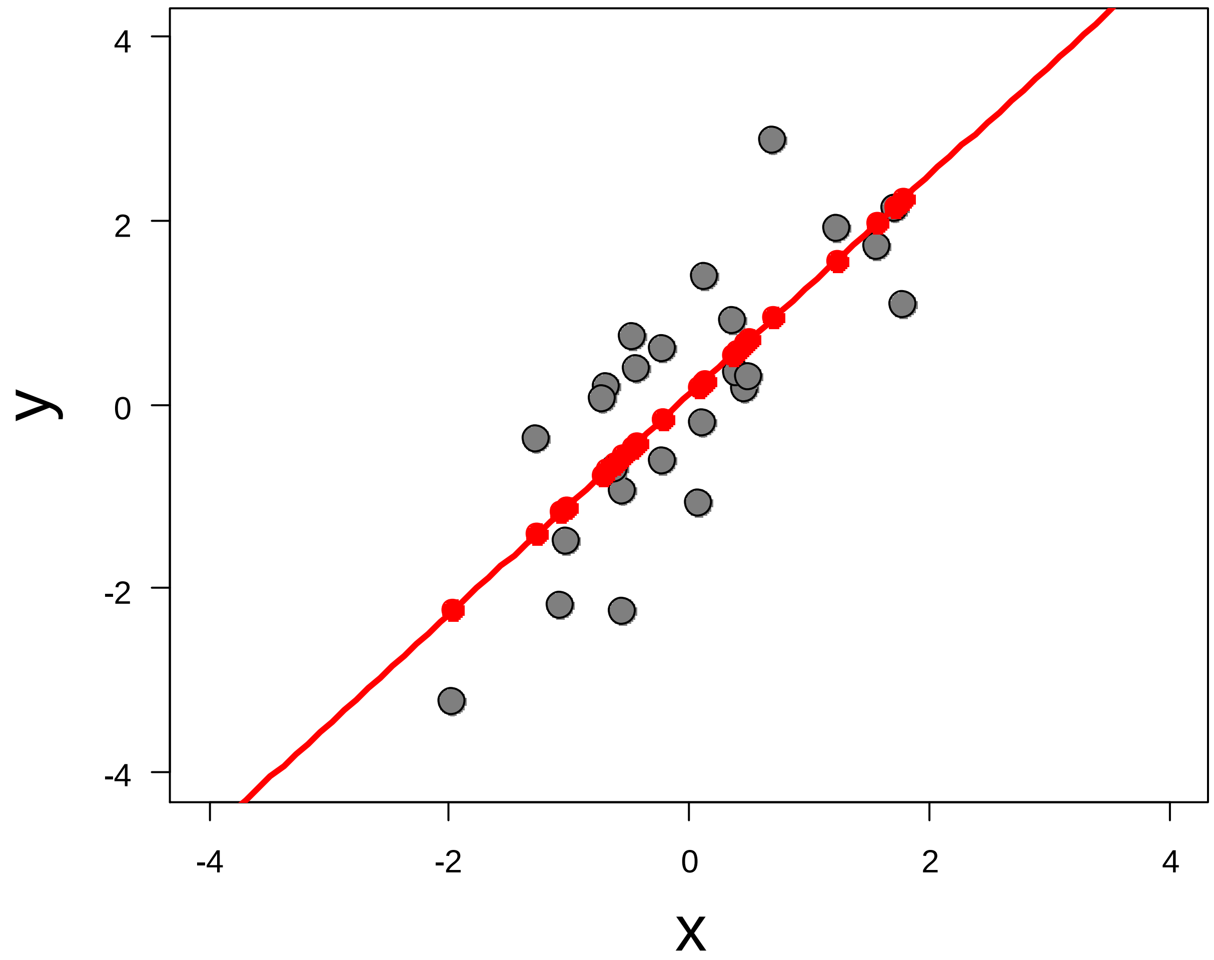
$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$

$\beta_0 = 0$

$\beta_1 = 1$

$\sigma = \sigma^2 = 1$

E(y)?



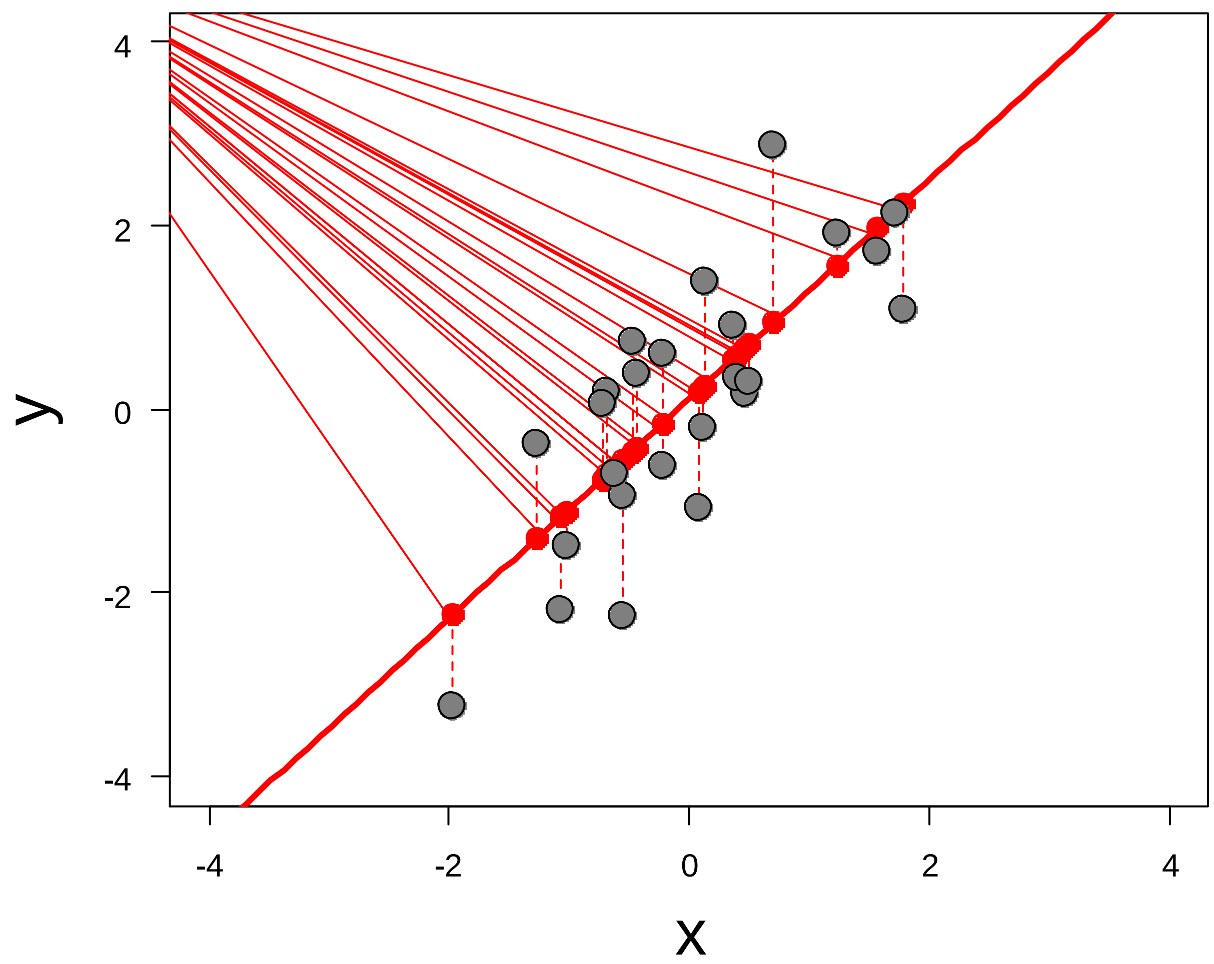
Data simulation

$$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$$
$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$
$$\beta_0 = 0$$
$$\beta_1 = 1$$
$$\sigma = \sigma^2 = 1$$

Expected values

$$\boldsymbol{E}(\boldsymbol{y}) = \boldsymbol{\eta} = \hat{\beta}_0 + \hat{\beta}_1 \boldsymbol{x}$$
$$\hat{\beta}_0 = 0.108$$
$$\hat{\beta}_1 = 1.194$$

What are residuals?



Data simulation

$$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

$$\beta_0 = 0$$

$$\beta_1 = 1$$

$$\sigma = \sigma^2 = 1$$

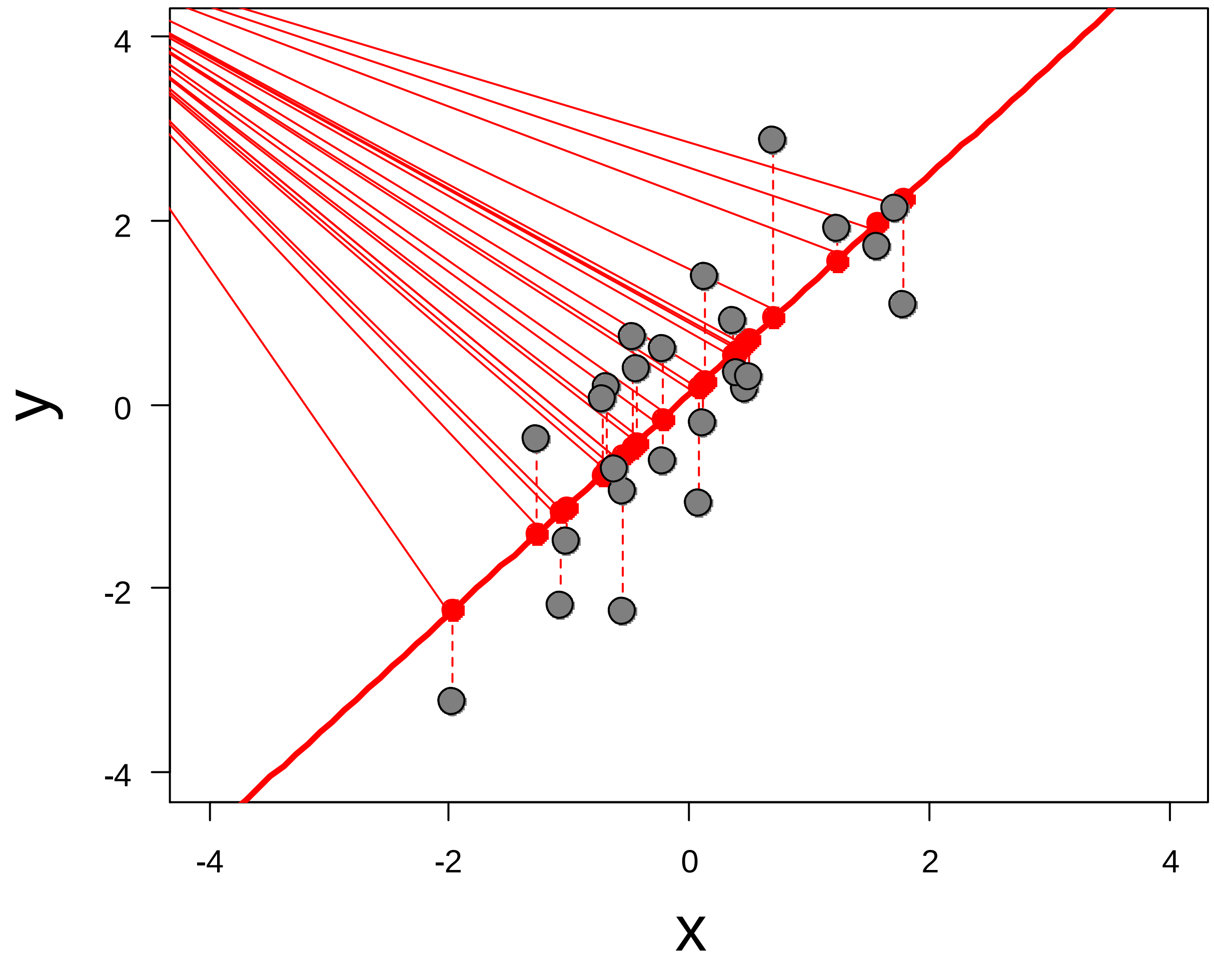
Expected values

$$\boldsymbol{E}(\boldsymbol{y}) = \boldsymbol{\eta}$$

Residuals

$$\boldsymbol{y} - \boldsymbol{\eta}$$

Residual sum of squares (RSS)



Residual sum of squares (RSS)

$$\mathrm{RSS} = \sum_{i=1}^{n} (y_i - \eta_i)^2$$

Add up the squared differences between data (y) and prediction (η) for every data point

As RSS increases, explanatory power declines…

AIC

AIC (general equation)

This is the log-likelihood (we work on the log-scale to prevent extreme values)

$$\mathrm{AIC} = 2k - 2 * \ln(\hat{L})$$

k is the number of parameters (this penalizes complex models)

AIC (for linear models)

$$\mathrm{AIC} = 2k + n \times \left(\ln \left(2 \times \pi \times \frac{\mathrm{RSS}}{n} \right) + 1 \right)$$

This is an alternative way to calculate AICc for linear models

AICc (the lil' c means corrected for small sample sizes)

AICc

$$\text{AICc} = \text{AIC} + \boxed{\frac{2k^2 + 2k}{n - k - 1}}$$

This is an extra penalty for complex models with small sample sizes

If n = 10 and k = 2, then this adds up to 1.714 (a substantial penalty on the AICc scale)

If n = 1000 and k = 2, then this adds up to 0.012 (a tiny penalty on the AICc scale)

If n = 10000 and k = 2, then this adds up to 0.0012 (a teeny, weeny tiny penalty on the AICc scale)

1a. How does model selection work?

Uninformative Parameters and Model Selection Using Akaike's Information Criterion

TODD W. ARNOLD,[1] *Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, St. Paul, MN 55108, USA*

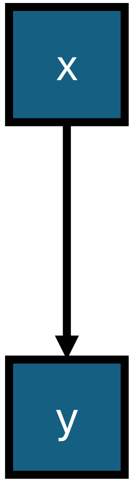
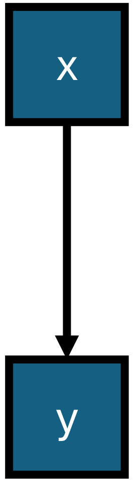


Table 1. Models examining effects of various covariates on detection probabilities of indicated breeding pairs of waterfowl in North Dakota, USA (from Pagano and Arnold 2009). I added single parameters assuming an additive effect to the base model, which included $K = 16$ parameters (8 species × 2 observers). Three of these covariates were considered biologically feasible (total ducks, vegetative cover, and cover type), 6 were not (random 5, 4, 8, and 1; not Sunday, Monday, or Wednesday; and last duck seen a mallard), and I excluded 6 additional nonsense or random variables (ΔAIC = 0.64–2.00) from presentation. I evaluated all models compared to the base model using Akaike's Information Criterion (AIC), ΔAIC, and changes in model deviance (Dev).

Model	AIC	ΔAIC	*K*	Dev
Total ducks	4,426.71	−16.62	17	4,392.71
Random 5	4,439.95	−3.38	17	4,405.95
Random 4	4,442.20	−1.13	17	4,408.19
Vegetative cover	4,442.65	−0.68	17	4,408.65
Random 8	4,442.81	−0.52	17	4,408.80
Random 1	4,442.90	−0.43	17	4,408.90
Base model	4,443.34	0	16	4,411.33
Not Sunday, Monday, or Wednesday	4,445.09	1.75	17	4,411.08
Cover type	4,445.25	1.92	17	4,411.25
Last duck seen a mallard	4,445.33	1.99	17	4,411.32

Arnold **(2010)** *Journal of Wildlife Management*

Simulate some data

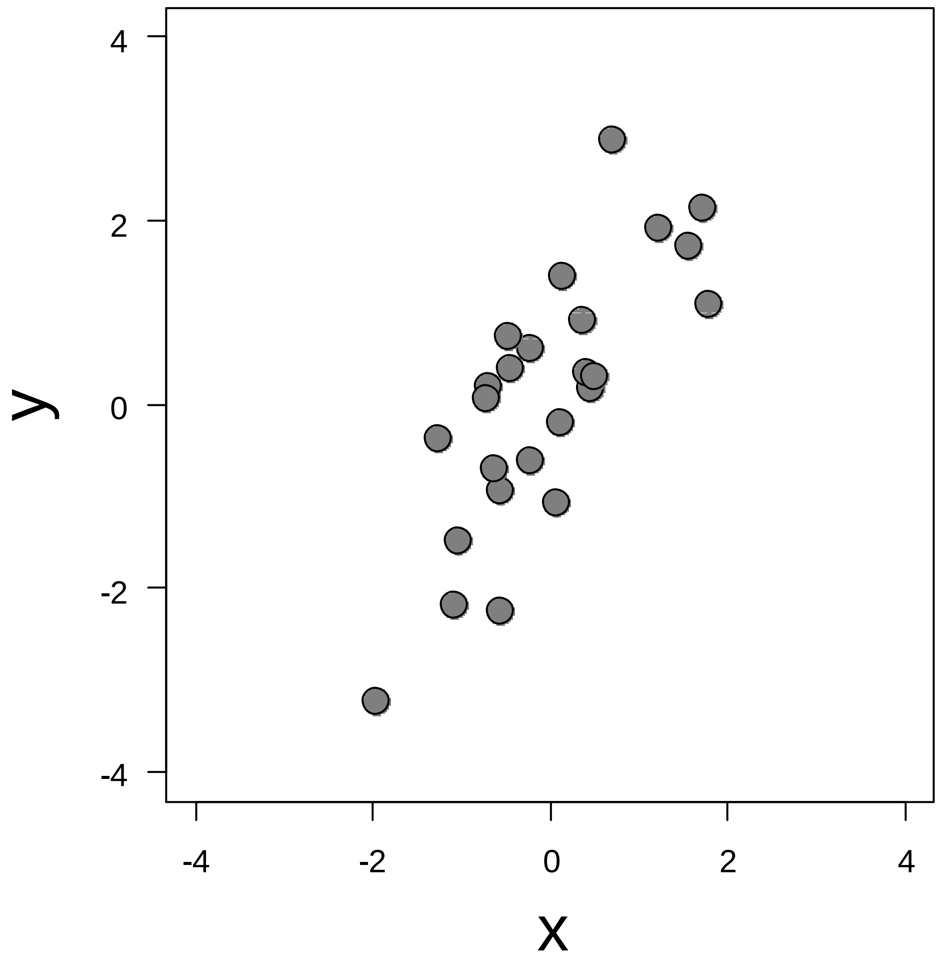


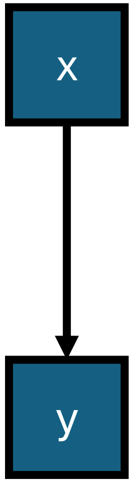
$$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{x}, 1)$$

$$\mathrm{n} = 25$$

Let's do that once…





$$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{x}, 1)$$

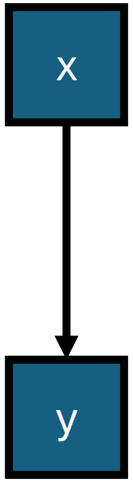
$$\mathrm{n} = 25$$

And compare two models

m0

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0, \sigma^2)$$

```
glm(y ~ 1)
```

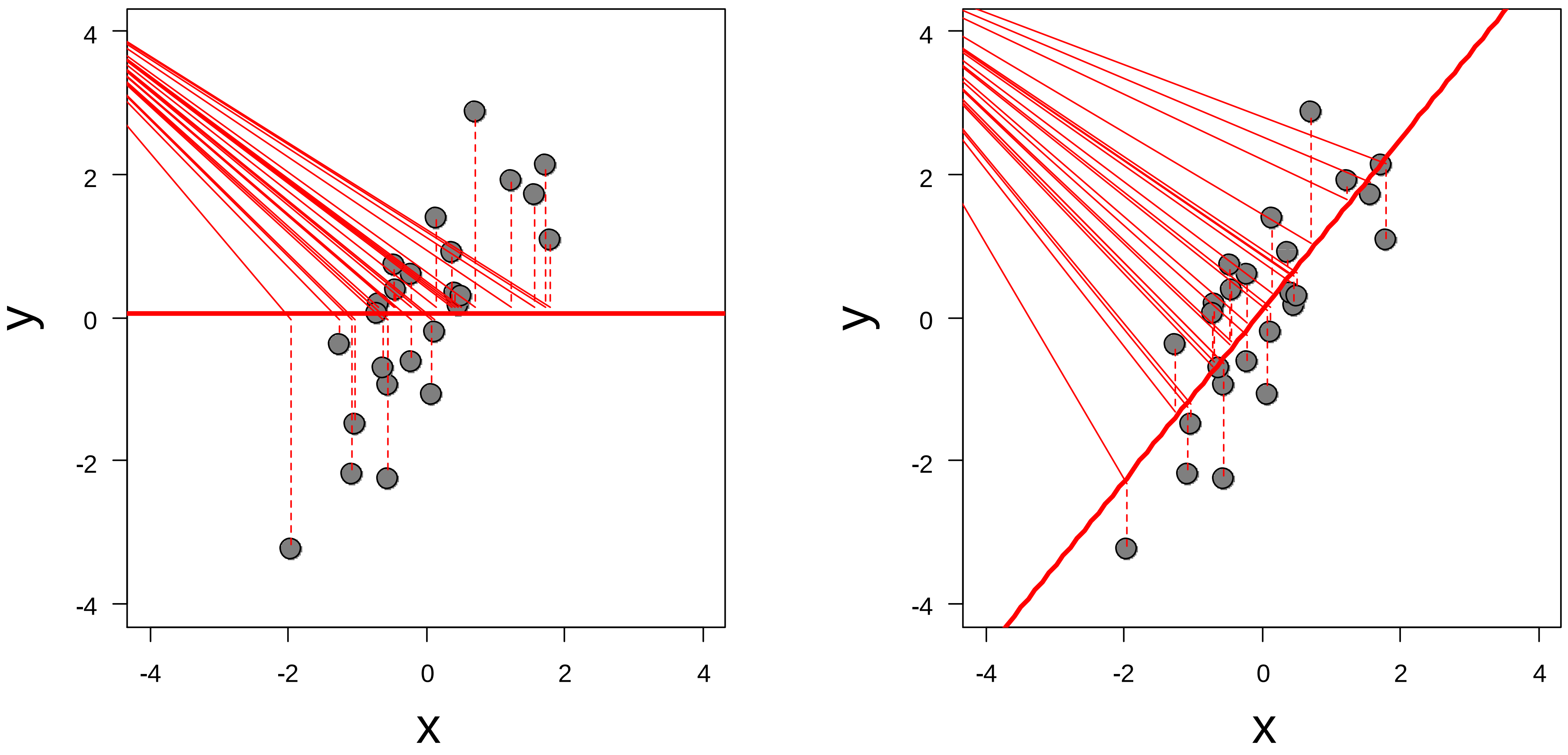


m1

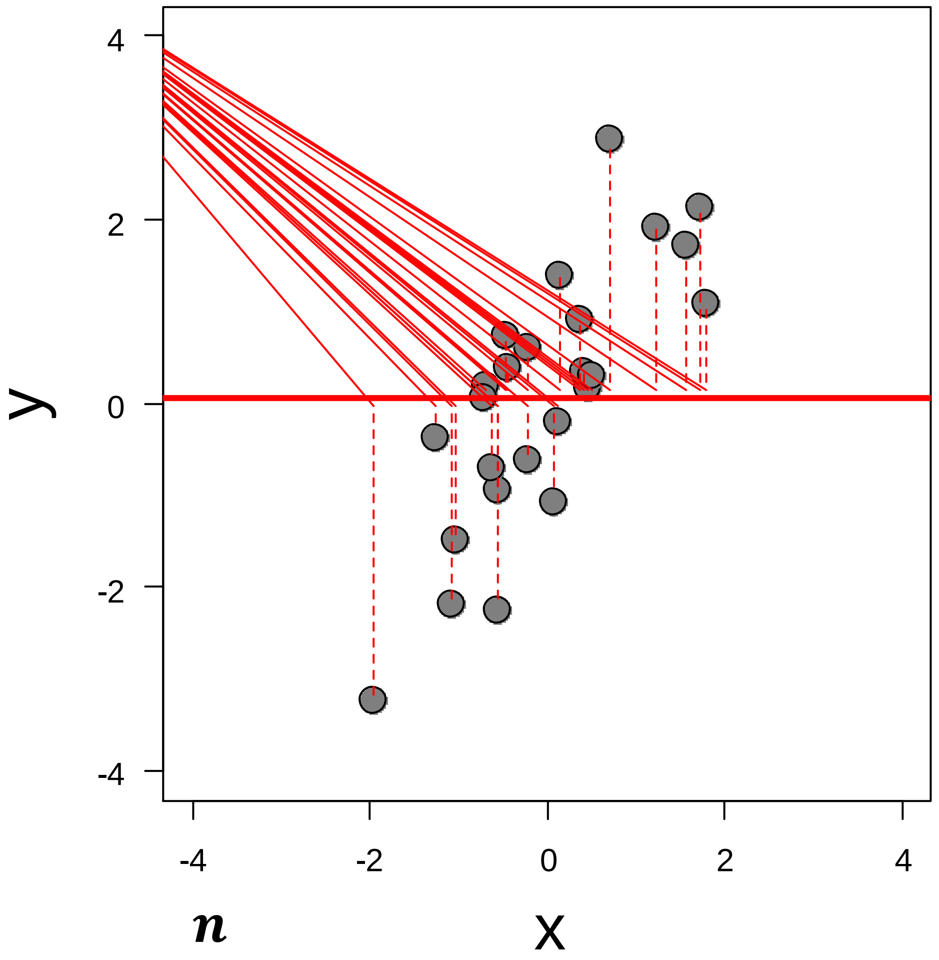
$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

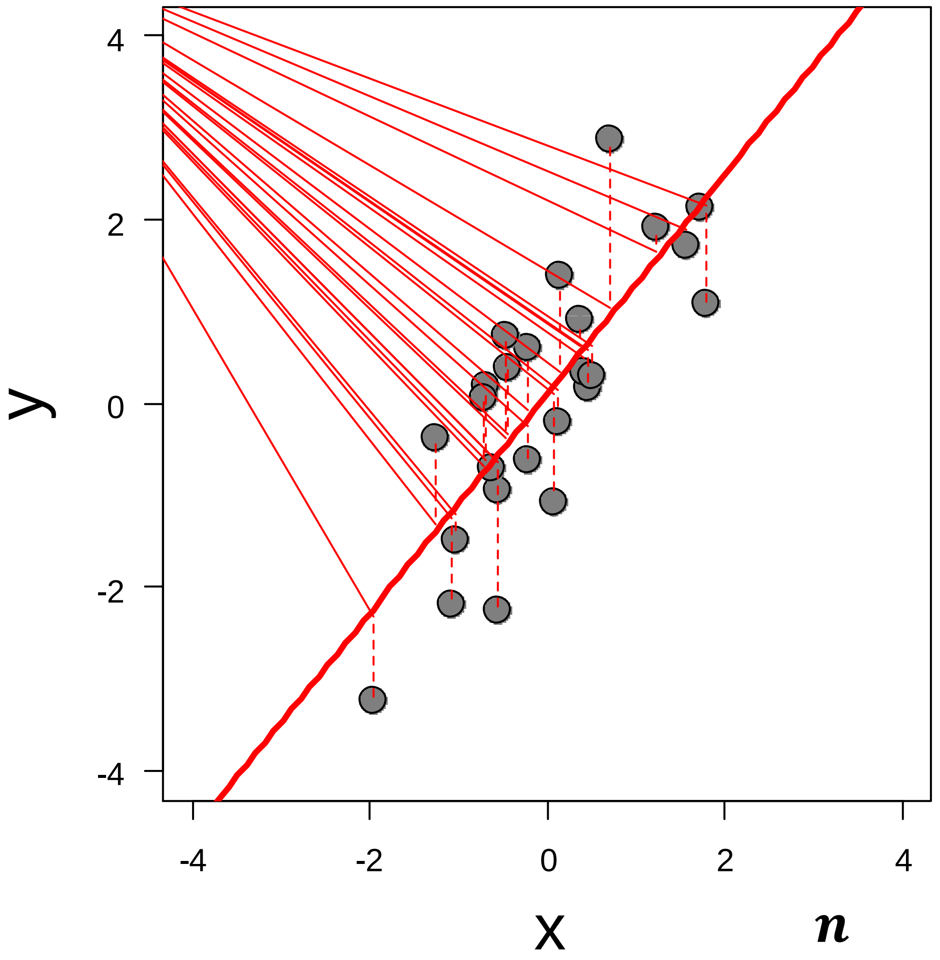
```
glm(y ~ x)
```

And compare two models from a single run



And compare two models from a single run (residuals)

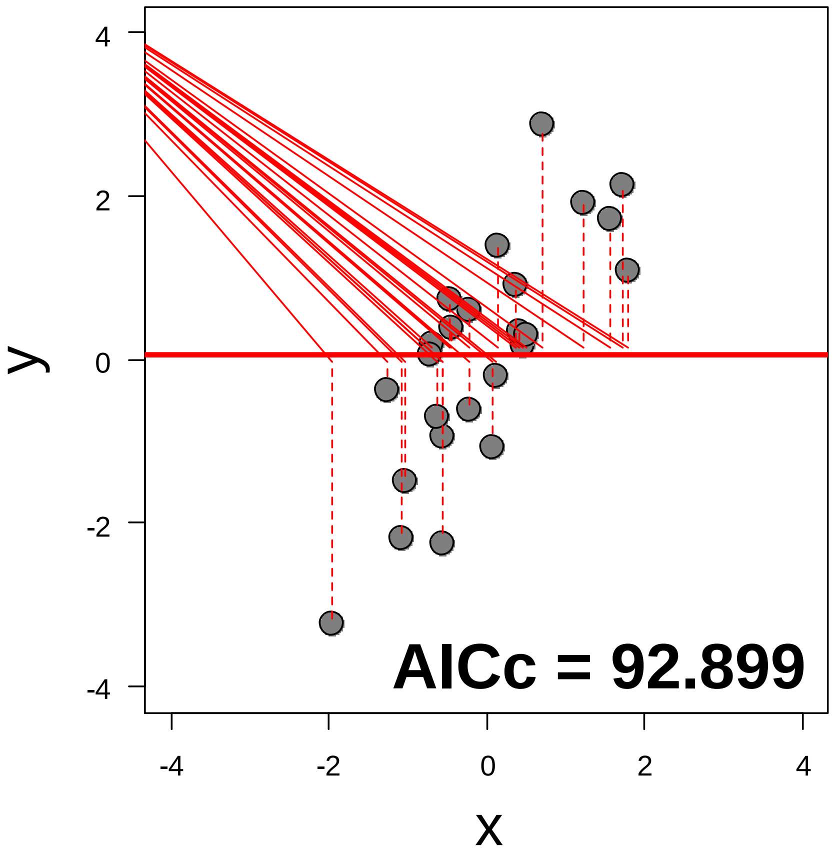


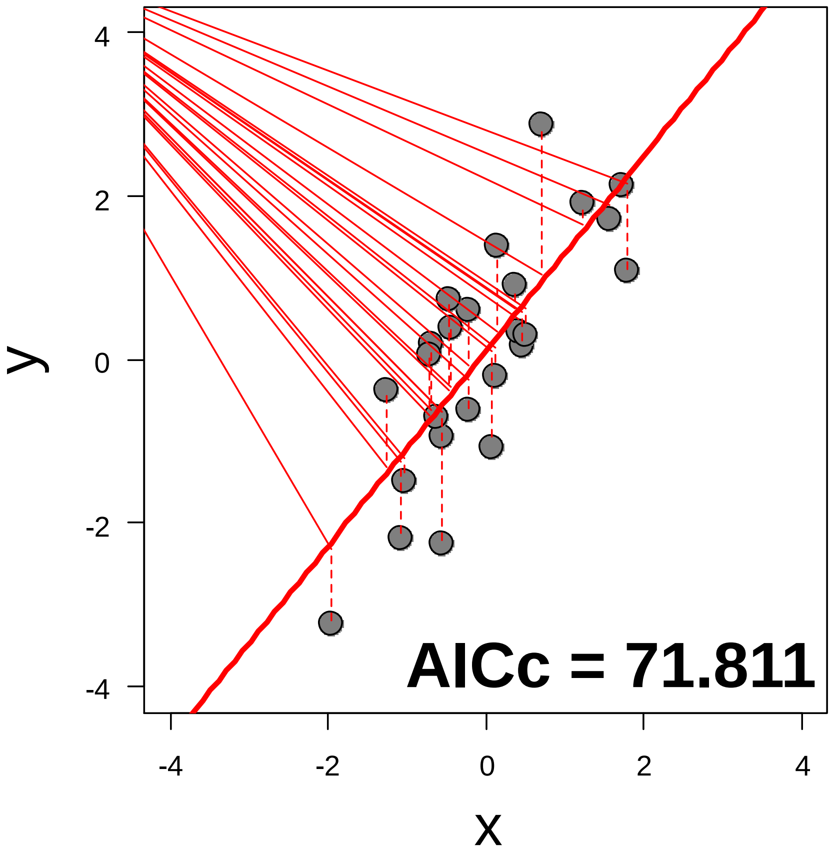


$$50.156 = \sum_{i=1}^{n} (y_i - \eta_i)^2$$

$$19.448 = \sum_{i=1}^{n} (y_i - \eta_i)^2$$

And compare two models from a single run (AICc)



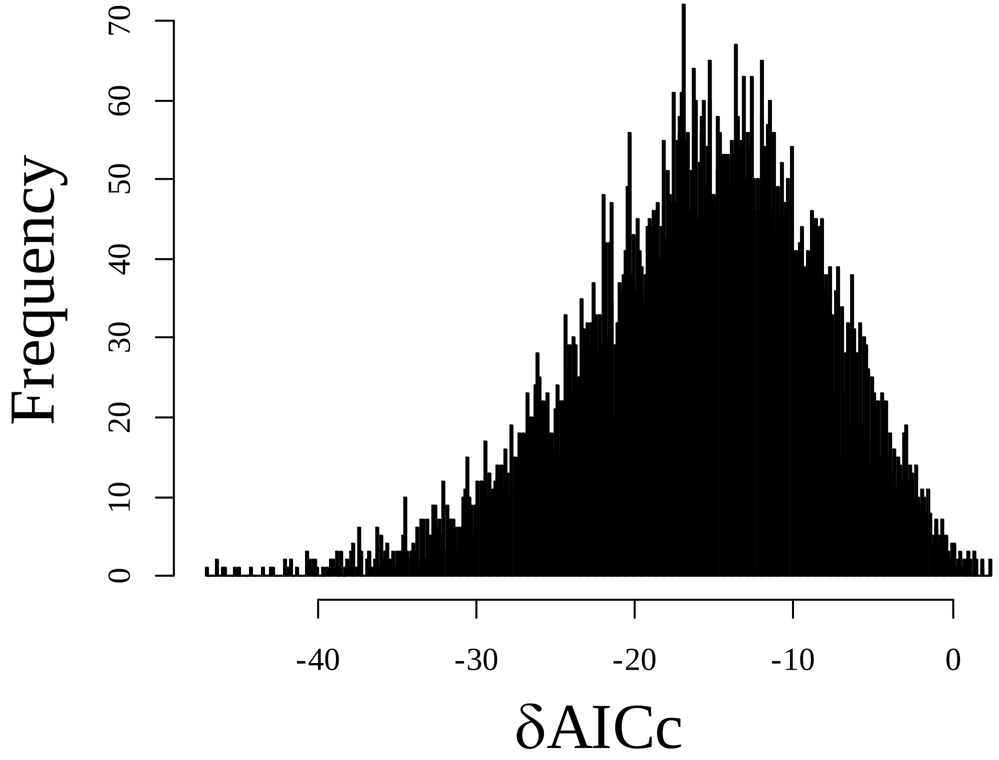


$$\mathrm{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\mathrm{RSS}}{n}\right) + 1\right) + \frac{2k^2+2k}{n-k-1}$$

Cool! Let’s do that 10,000 times!

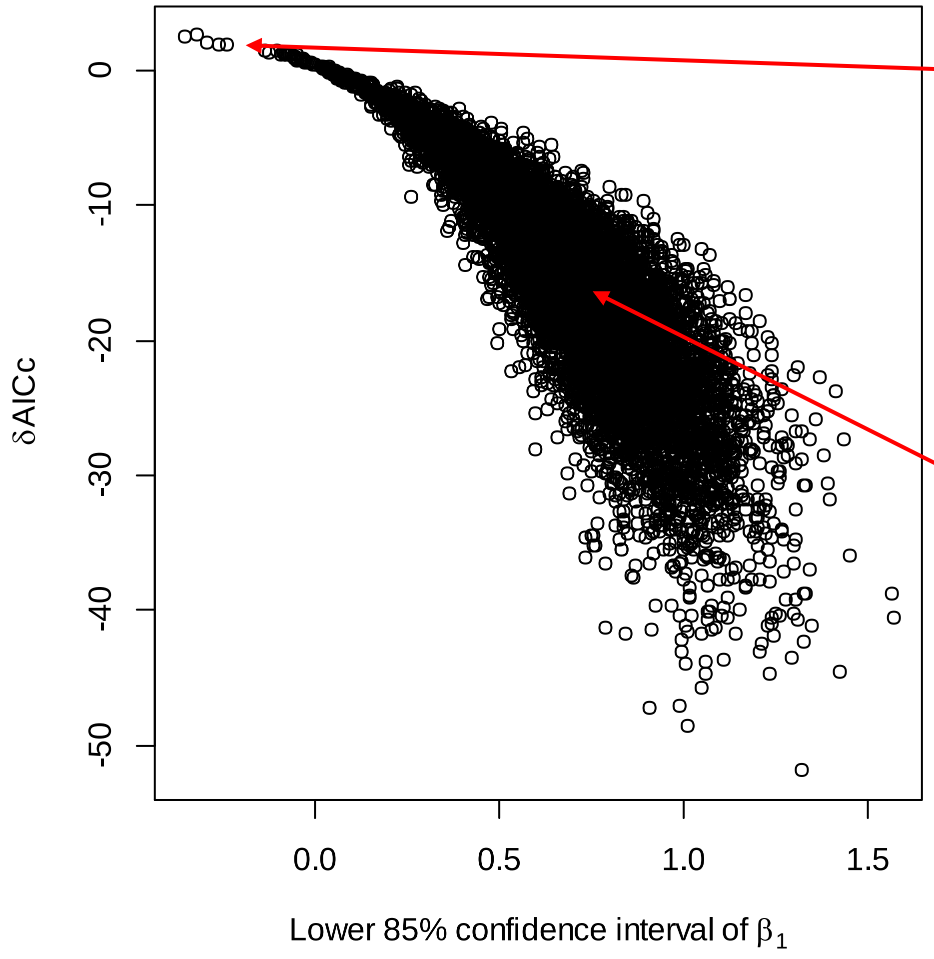
And compare two models from 10k runs

- mean(δAIC) = -15.89
- median(δAIC) = -15.38
- In 0.3% of simulations, m0 (the intercept only model) was superior to m1 (the true model).
- In 1.5% of simulations, m0 was ~equivalent to m1 (δAIC < 2).
- m0 had some support (δAIC < 7) in 11% of simulations.



The data generating model is consistently the 'best' model

Relationship between 85% confidence intervals and AIC/AICc



If we aren't confident x has an effect… then no improvement over simpler model

If we are confident that the effect of x is > 0, then the model with that covariate will have a lower AIC/AICc

$$\text{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\text{RSS}}{n}\right) + 1\right) + \frac{2k^2 + 2k}{n-k-1}$$

So, what are our take-homes from ‘example 1a’

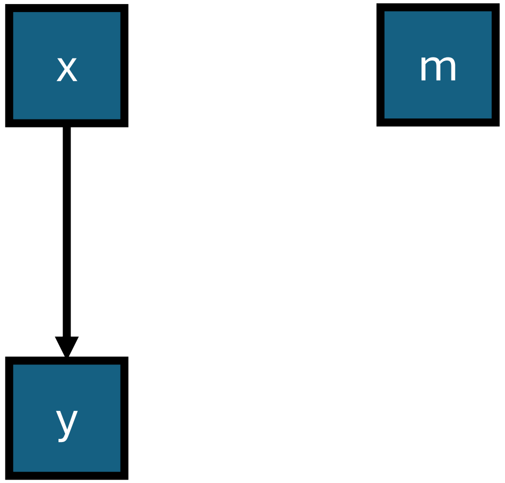
- **Including a parameter that has a significant (85% CI) association with the response will lead to lower AIC values (because more variance is explained)**

- **In this case, it was causal. That isn’t necessarily true!**

1b. How does model selection work?

let’s break some stuff!

1b. Let's add some meaningless information (*m*)



Uninformative Parameters and Model Selection Using Akaike's Information Criterion

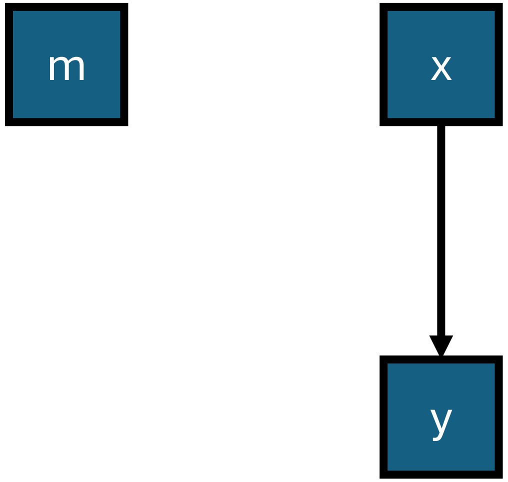
TODD W. ARNOLD,[1] *Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, St. Paul, MN 55108, USA*

Table 1. Models examining effects of various covariates on detection probabilities of indicated breeding pairs of waterfowl in North Dakota, USA (from Pagano and Arnold 2009). I added single parameters assuming an additive effect to the base model, which included $K = 16$ parameters (8 species × 2 observers). Three of these covariates were considered biologically feasible (total ducks, vegetative cover, and cover type), 6 were not (random 5, 4, 8, and 1; not Sunday, Monday, or Wednesday; and last duck seen a mallard), and I excluded 6 additional nonsense or random variables (ΔAIC = 0.64–2.00) from presentation. I evaluated all models compared to the base model using Akaike's Information Criterion (AIC), ΔAIC, and changes in model deviance (Dev).

Model	AIC	ΔAIC	*K*	Dev
Total ducks	4,426.71	−16.62	17	4,392.71
Random 5	4,439.95	−3.38	17	4,405.95
Random 4	4,442.20	−1.13	17	4,408.19
Vegetative cover	4,442.65	−0.68	17	4,408.65
Random 8	4,442.81	−0.52	17	4,408.80
Random 1	4,442.90	−0.43	17	4,408.90
Base model	4,443.34	0	16	4,411.33
Not Sunday, Monday, or Wednesday	4,445.09	1.75	17	4,411.08
Cover type	4,445.25	1.92	17	4,411.25
Last duck seen a mallard	4,445.33	1.99	17	4,411.32

Arnold **(2010)** *Journal of Wildlife Management*

Simulate some data



$$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{x}, 1)$$

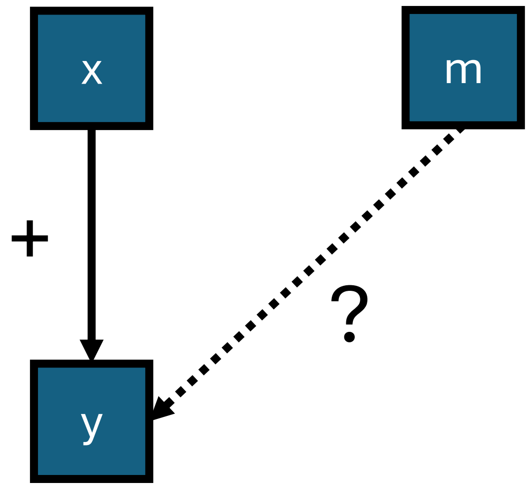
$$\mathrm{n} = 25$$

$$\boldsymbol{m} \sim \mathrm{Normal}(0,1)$$

1b. Similarly, we'll compare two models*

m1

$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$



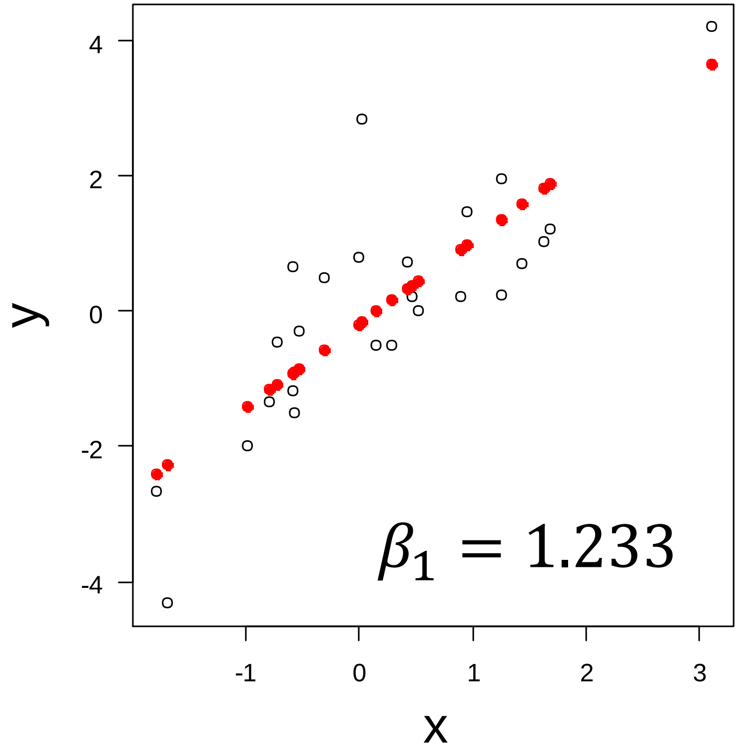
m2

$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$

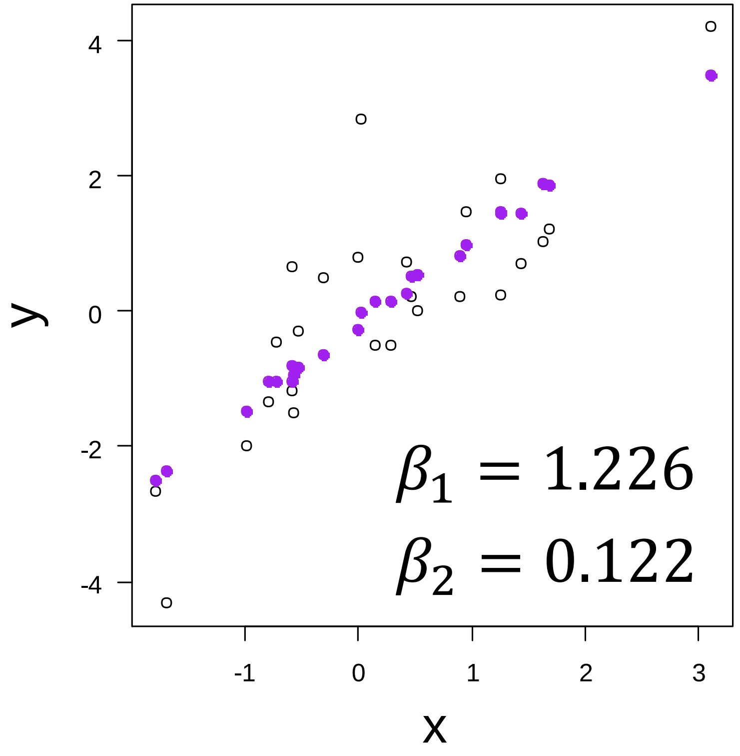
***Note that $\boldsymbol{y} \sim \mathrm{Normal}(\beta_0, \sigma^2)$ generally won't be competitive as in 1a**

1b. Predictions from both models

$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

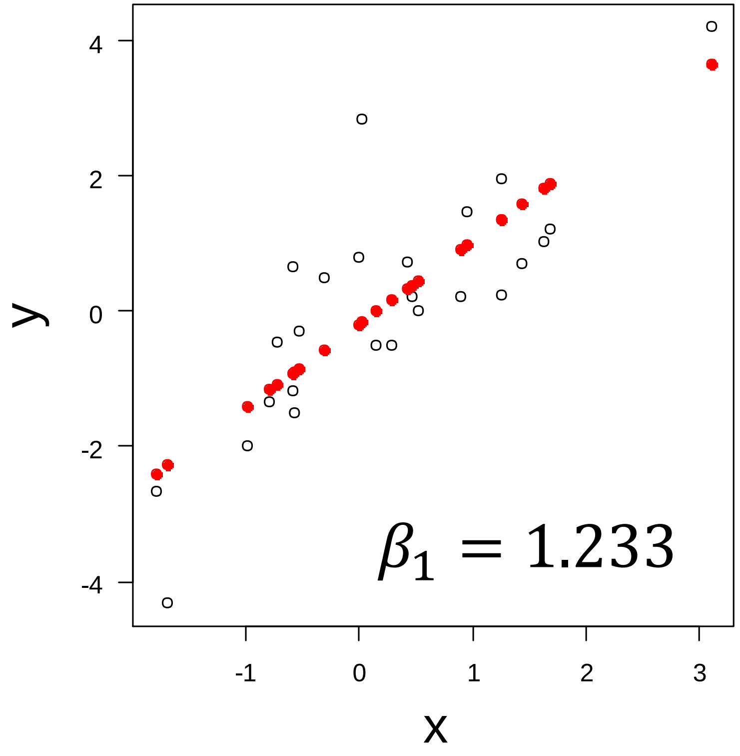


$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$



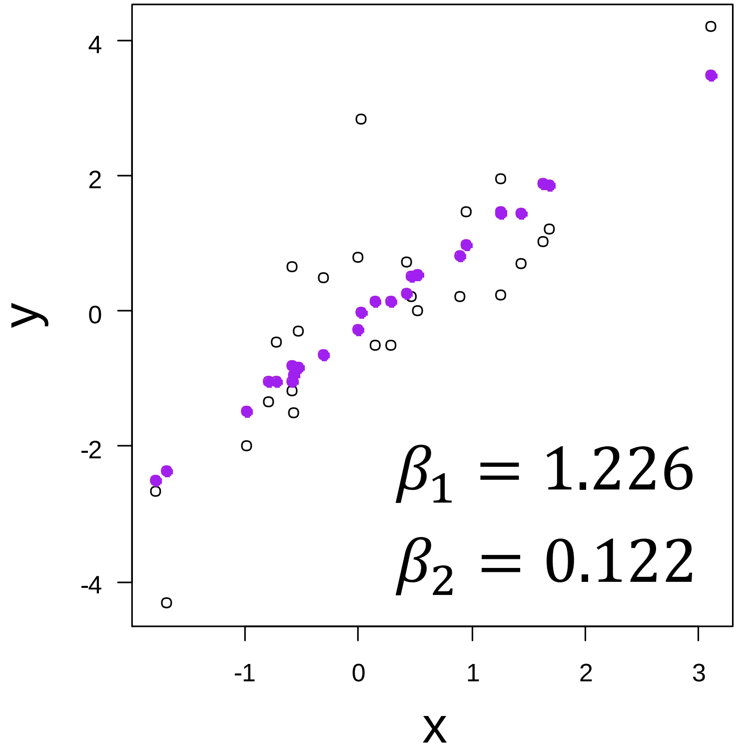
1b. RSS

$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$



$$24.839 = \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} (y_i - \eta_i)^2$$

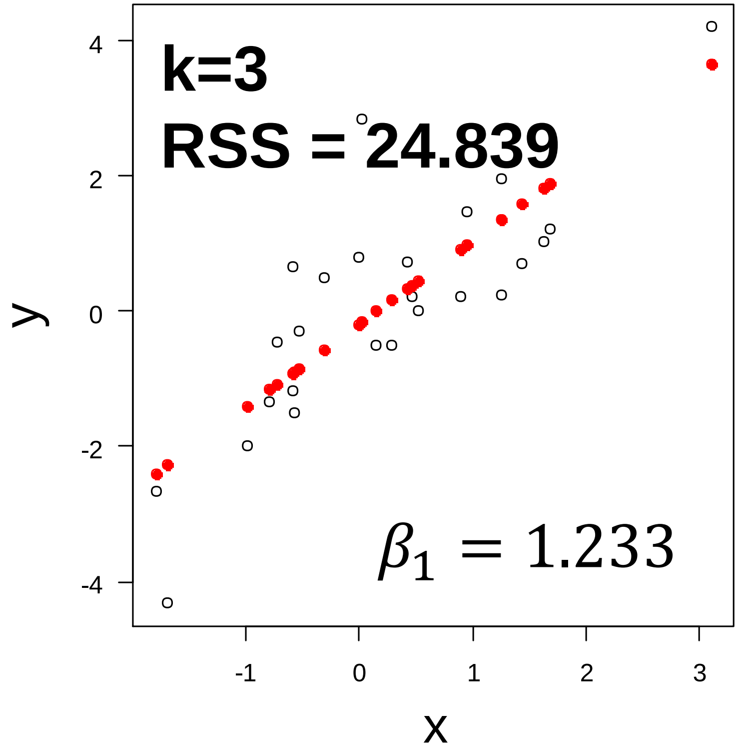
$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$



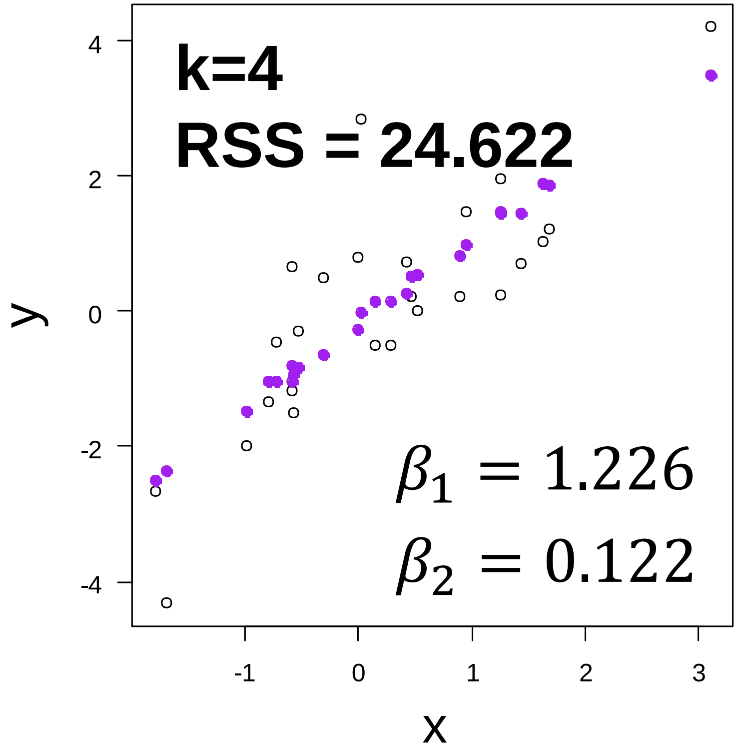
$$24.622 = \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} (y_i - \eta_i)^2$$

1b. What are our predictions for AIC

$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$



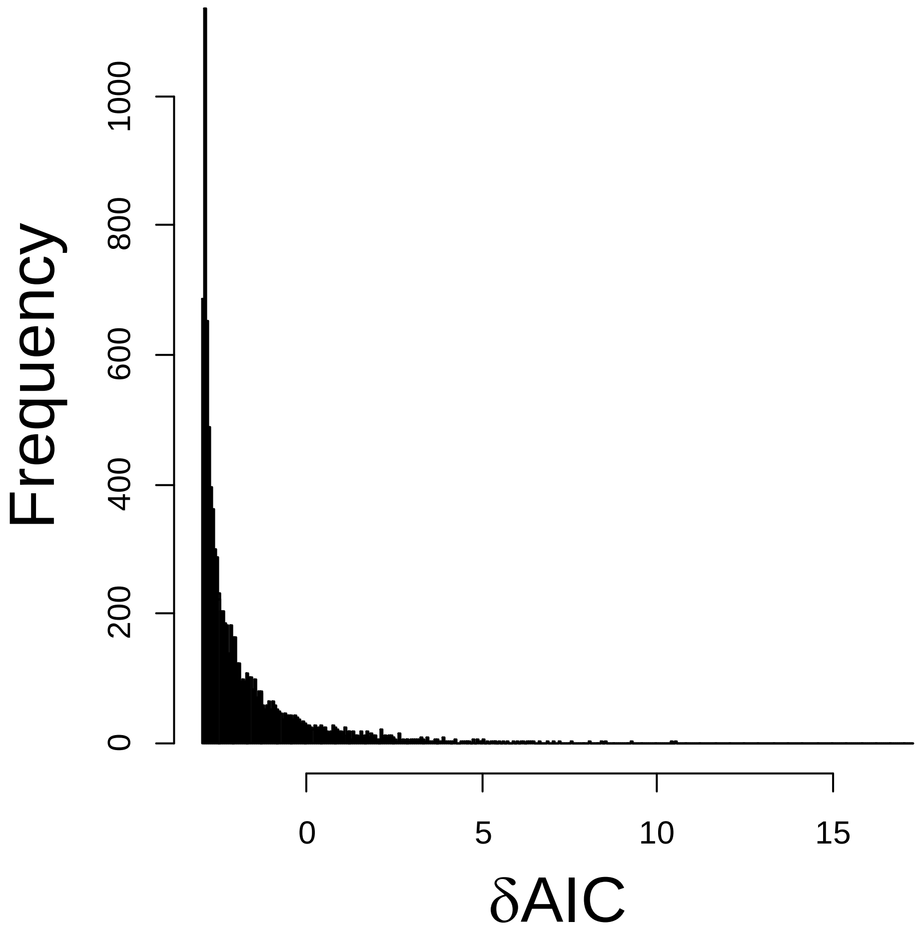
$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$



$$\text{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\text{RSS}}{n}\right) + 1\right) + \frac{2k^2 + 2k}{n - k - 1}$$

And compare two models from 10k runs

- mean(δAIC) = -1.704
- median(δAIC) = -2.316
- In 11.4% of simulations, m2 (the two covariate model) was superior to m1 (the true model).
- In 38.5% of simulations, m2 was ~equivalent to m1 (δAIC < 2).
- **m0 had some support (δAIC < 4) in 100% of simulations**.

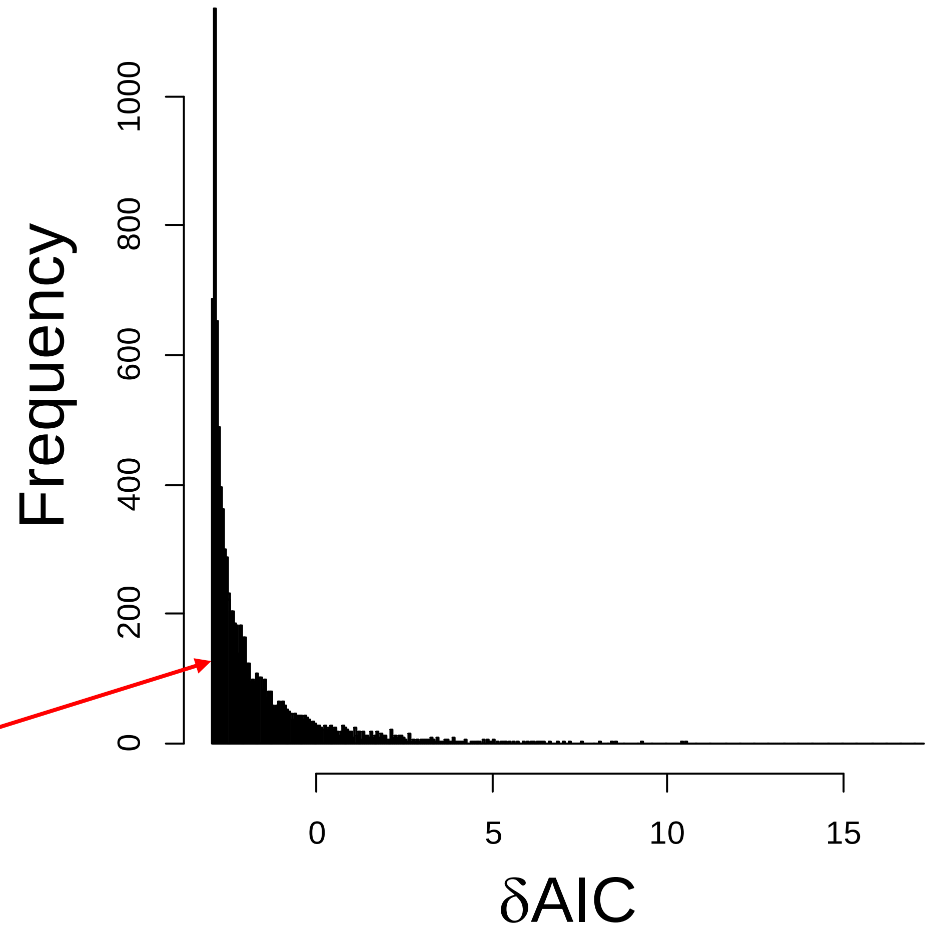


<u>Did it?!</u>

What’s happening here?

- **m0 had some support (δAIC < 4) in 100% of simulations.**

- **Did it?!**



Cual es el problema?

$$\mathrm{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\mathrm{RSS}}{n}\right) + 1\right) + \frac{2k^2+2k}{n-k-1}$$

What factors affect m2 (model including m) being 'important'?

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

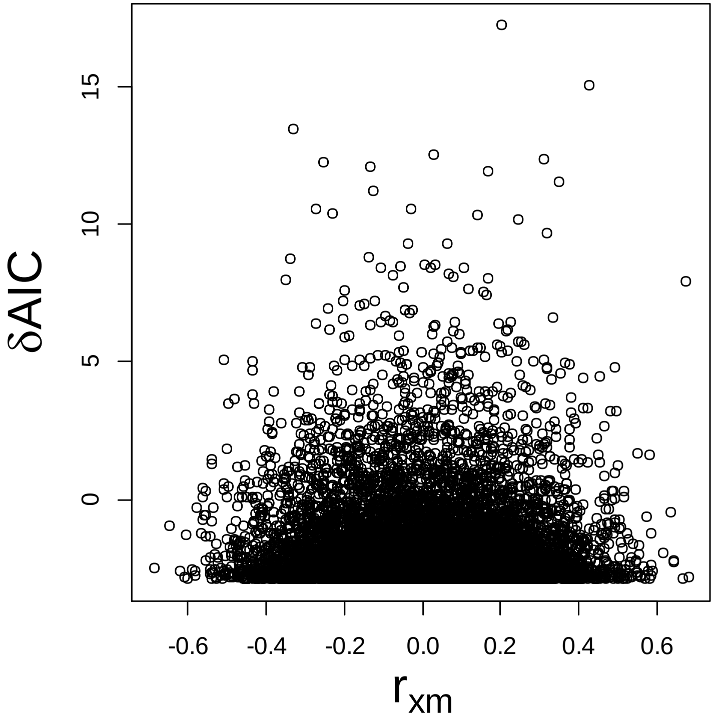
$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$

$$\mathrm{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\mathrm{RSS}}{n}\right) + 1\right) + \frac{2k^2 + 2k}{n - k - 1}$$

What factors affect m2 (model including m) being 'important'?

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$

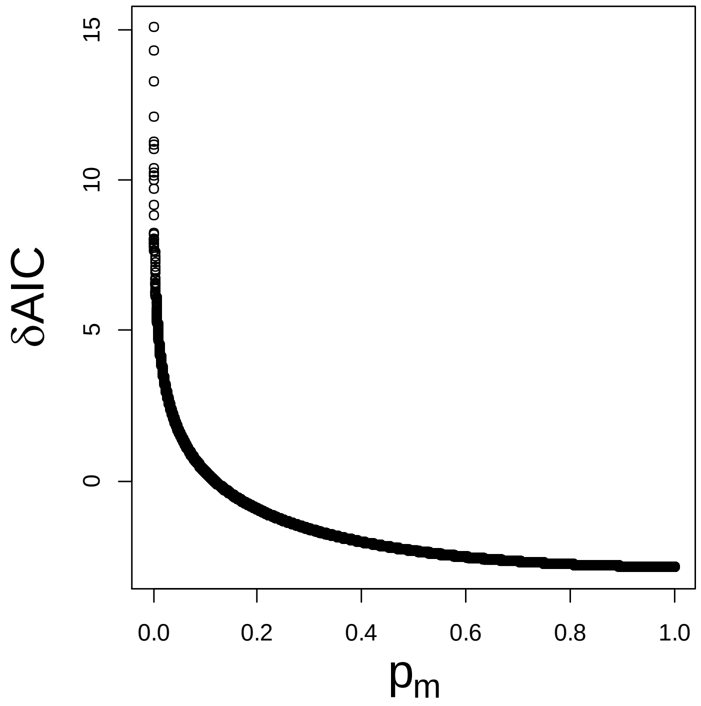


$$\mathrm{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\mathrm{RSS}}{n}\right) + 1\right) + \frac{2k^2 + 2k}{n - k - 1}$$

What factors affect m2 (model including m) being 'important'?

$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

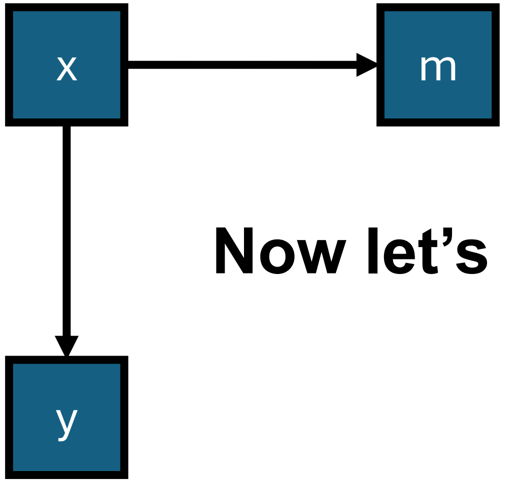
$$\boldsymbol{y} \sim \text{Normal}(\beta_0 + \beta_1 \boldsymbol{x} + \beta_2 \boldsymbol{m}, \sigma^2)$$



The effect of m had a p-value < 0.1 in approximately 10% of simulations (that makes sense!)

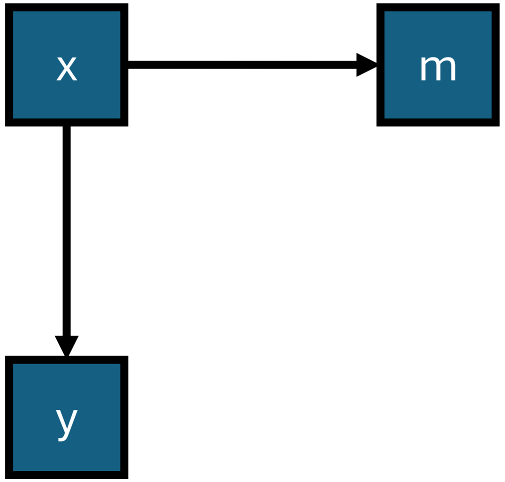
δAIC favored the two covariate model 11.4% of the time…

$$\text{AICc} = 2k + n \times \left(\ln\left(2 \times \pi \times \frac{\text{RSS}}{n}\right) + 1\right) + \frac{2k^2+2k}{n-k-1}$$



Now let's really break some stuff (10k times)!

How do complex models affect model selection? A simple simulated example



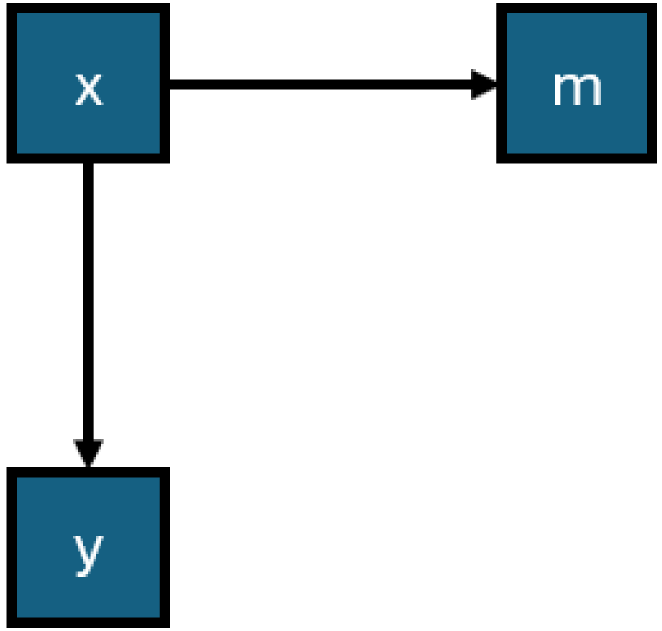
$$\boldsymbol{x} \sim \text{Normal}(0,1)$$

$$\boldsymbol{y} \sim \text{Normal}(\boldsymbol{x}, 1)$$

$$\boldsymbol{m} \sim \text{Normal}(\boldsymbol{x}, 1)$$

$$\text{n} = 25$$

Now we'll imagine that we misunderstand the data-generating process (fun!)



$\boldsymbol{x} \sim \mathrm{Normal}(0,1)$

$\boldsymbol{y} \sim \mathrm{Normal}(\boldsymbol{x}, 1)$

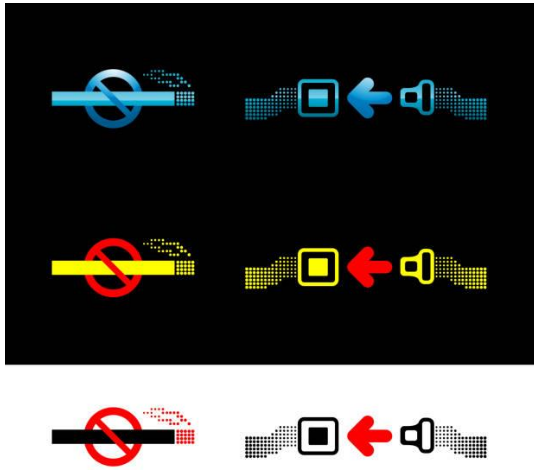
$\boldsymbol{m} \sim \mathrm{Normal}(\boldsymbol{x}, 1)$

$\mathrm{n} = 25$

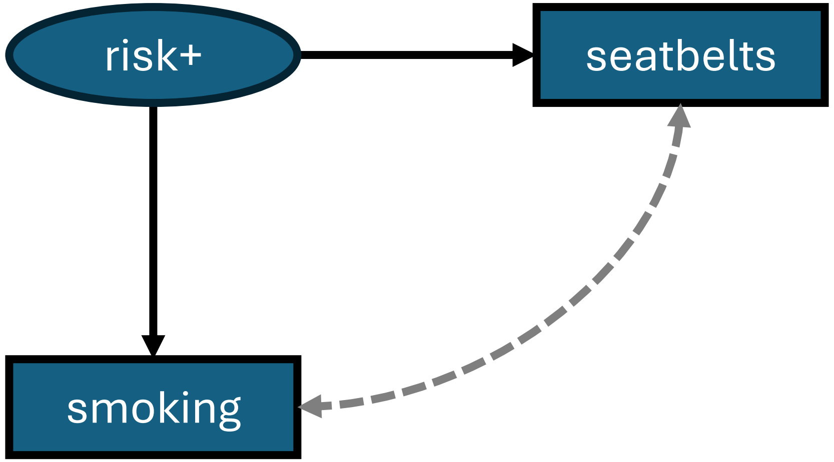
Let's assume that we think ***m*** affects **y**

This isn’t as wild as it seems

Eiser et al. (**1979**) *Addictive Behavior*
Grout et al. (**1983**) *Public Health*
Cliff et al. (**1982**) *Public Health*
Helsing and Comstock (**1977**) *American Journal of Public Health*
Manheimer et al. (**1966**) *Traffic Safety Research Review*
Williams (**1973**) *Journal of Health and Social Behavior*

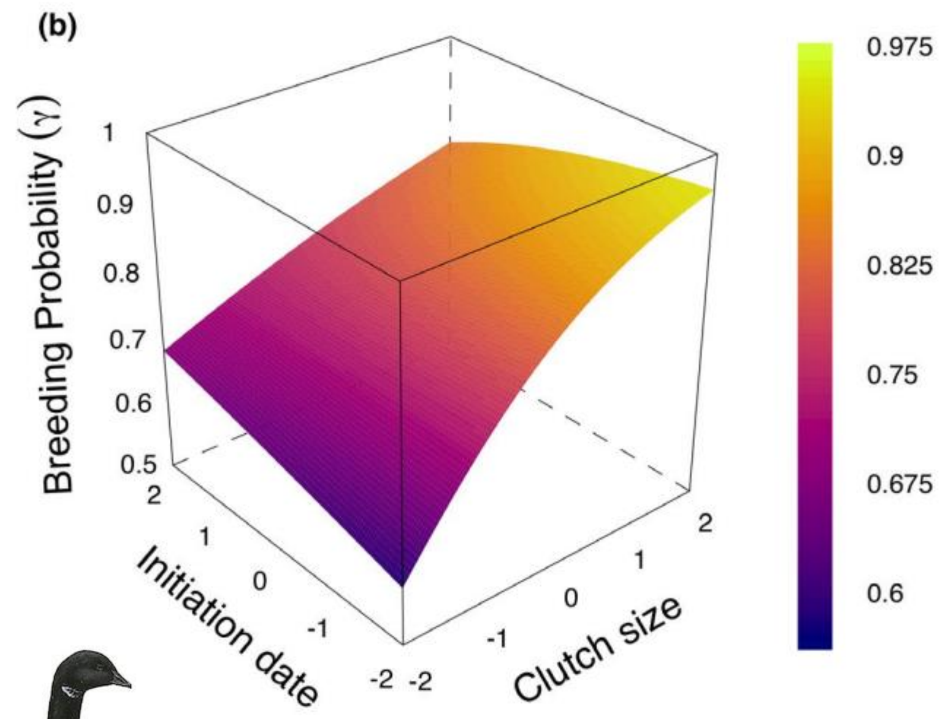


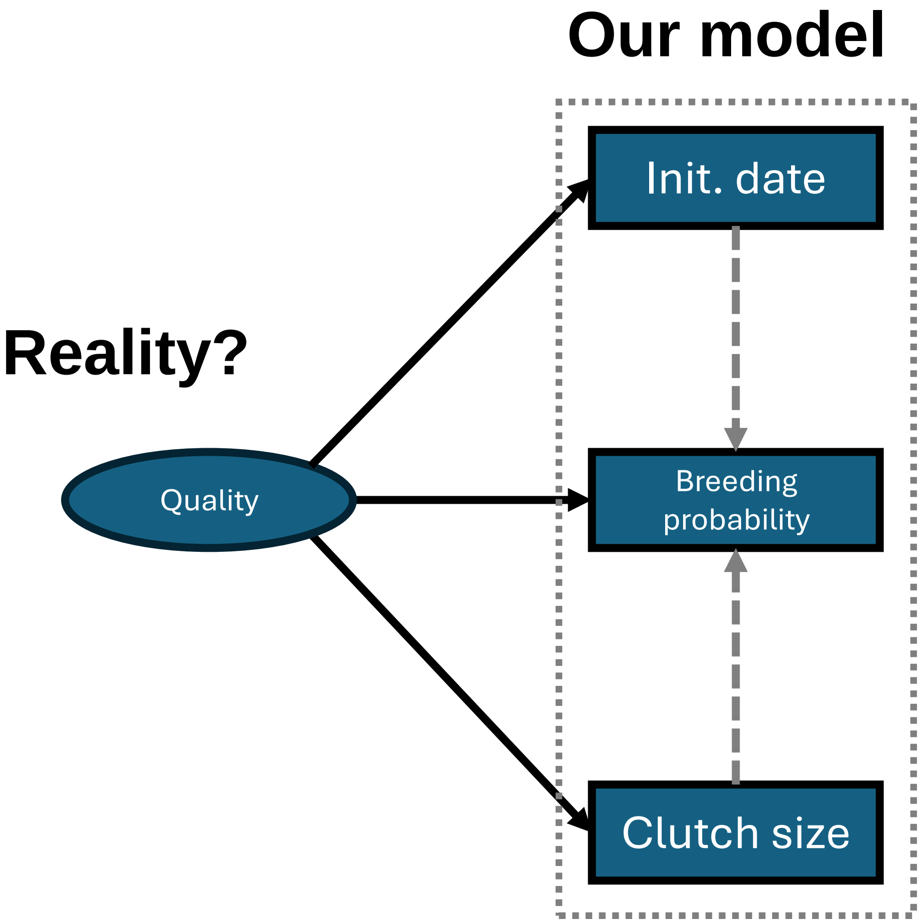
Indeed, in a study conducted in 2006 as part of a tobacco litigation, seat-belt usage was listed as one of the first variables to be controlled for.
-Pearl and Mackenzie (**2018**) *The Book of Why*



Smoking and seat-belt use

This isn't as wild as it seems

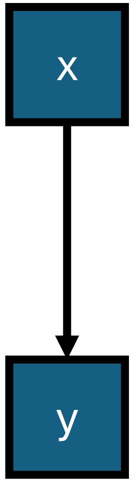




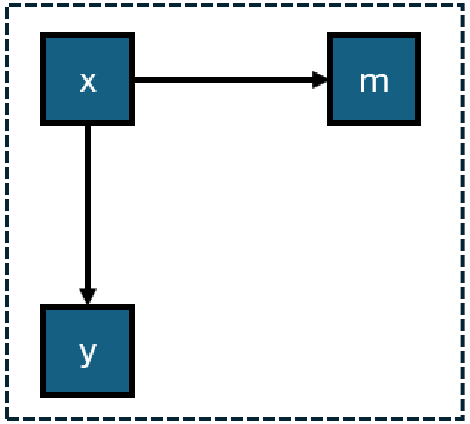
Lohman, Riecke, et al. **(2021)** *Ecology and Evolution*

Back to our simulation, we'll compare two models (power analysis)

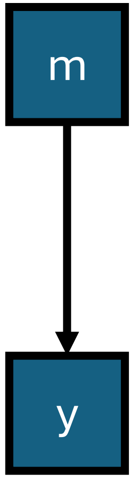
m1



Reality



m2



$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{x}, \sigma^2)$$

$$\boldsymbol{y} \sim \mathrm{Normal}(\beta_0 + \beta_1 \boldsymbol{m}, \sigma^2)$$

model_selection_and_model_weights.R

We'll do this 10k times (brute force to the rescue!) and save:

1. **The difference in AICc between the models (δAIC)**
   - Differences of < 2 indicate model 'equivalence'
   - Differences of < 7 indicate some support and should rarely be dismissed

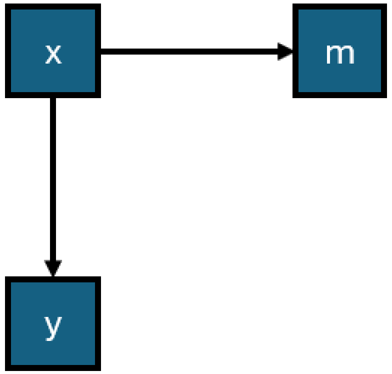
2. **The model weights (a function of δAIC that tells us the relative likelihood of each model)**

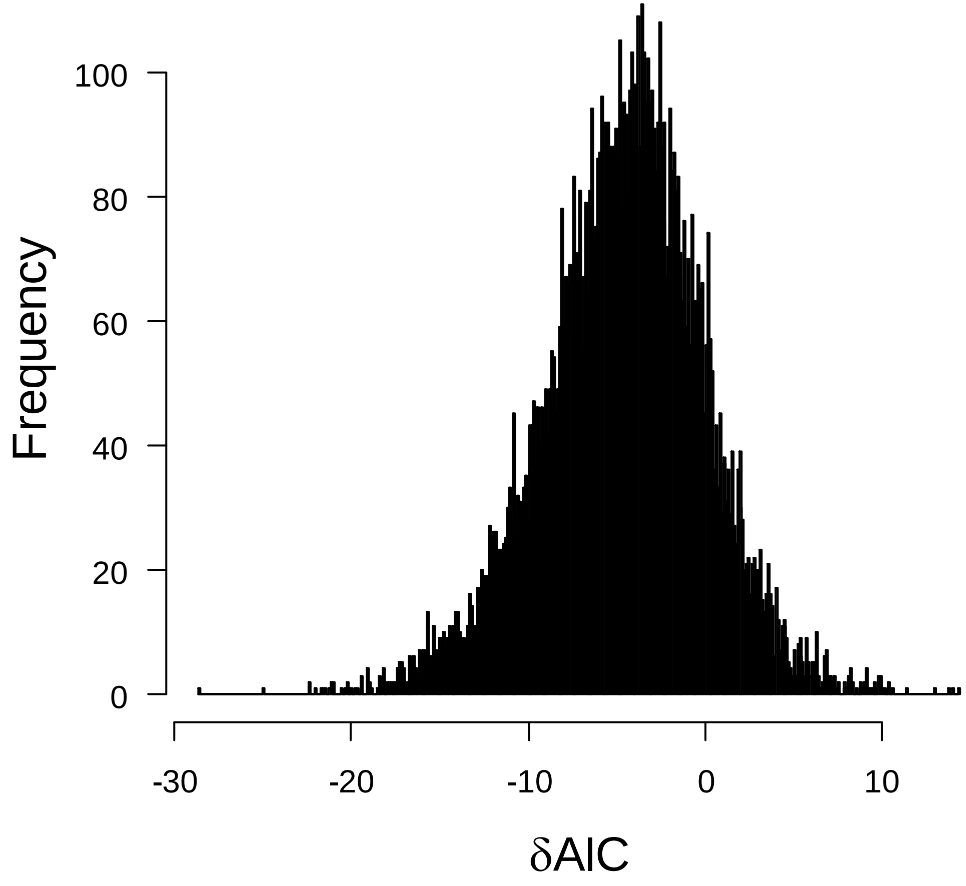
3. **The observed correlation between *x* and *m***

Burnham, Anderson, and Huyvaert (**2011**) *Behavioral Ecology and Sociobiology*

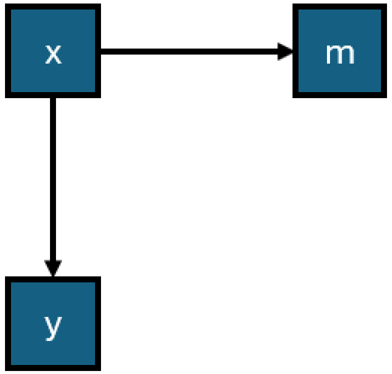
Results (n = 25; 10k sims)

- mean(δAIC) = -4.57
- median(δAIC) = -4.35
- In 14% of simulations, m2 (the spurious model) was superior to m1 (the true model).
- In 28% of simulations, m2 was ~equivalent to m1 (δAIC < 2).
- m2 had some support (δAIC < 7) in 72% of simulations.

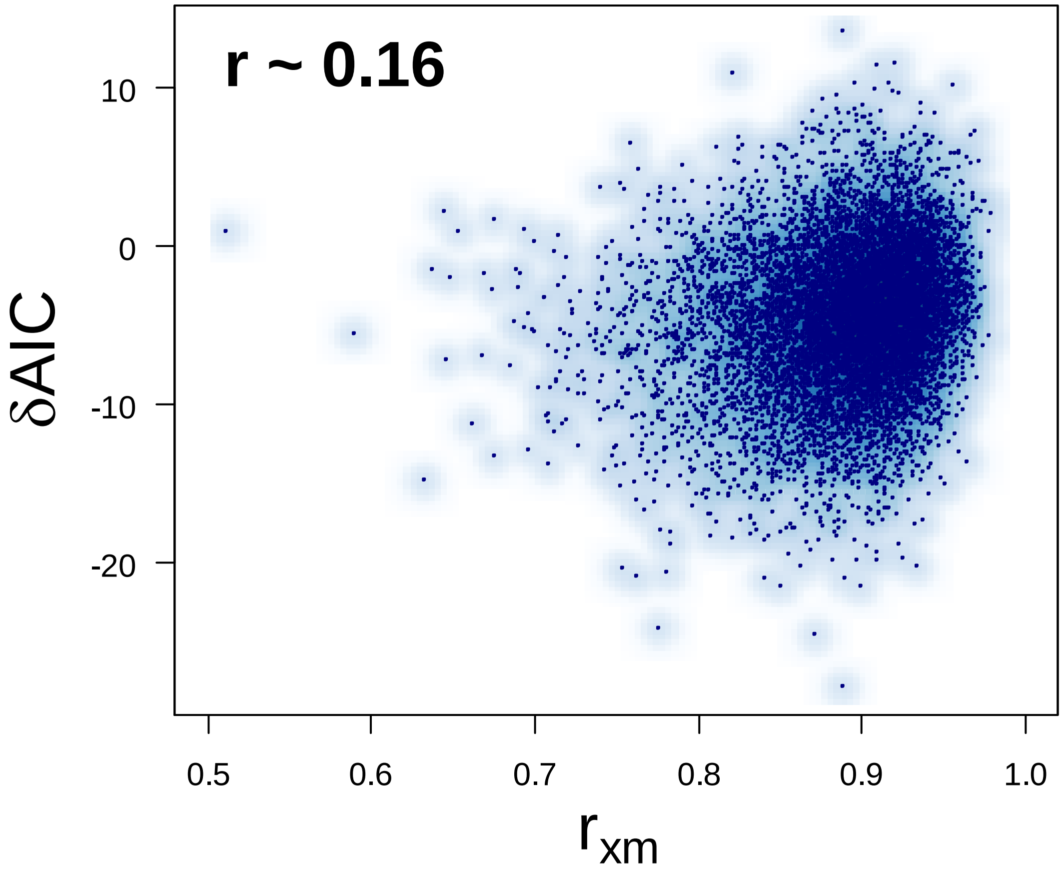




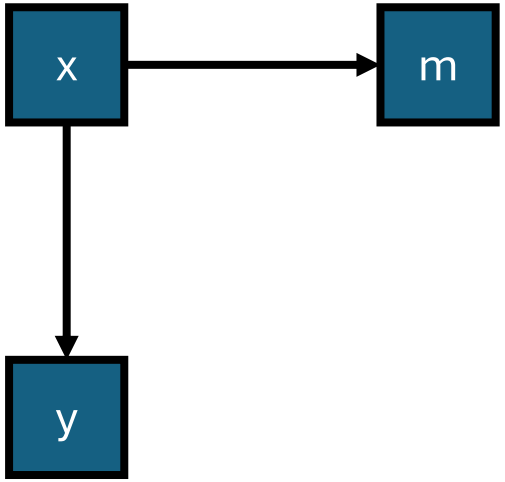
Results (n = 25; 10k sims)



- Increases in correlations among x and m led to difficulty in resolving differences among competing models



What would happen if we changed *n*?



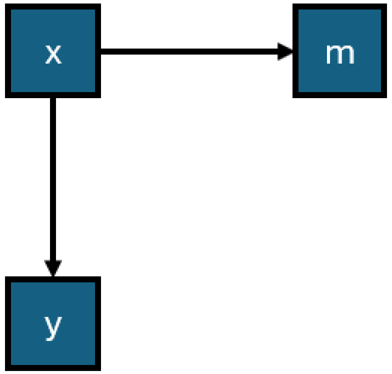
$$\boldsymbol{x} \sim \text{Normal}(0,1)$$

$$\boldsymbol{y} \sim \text{Normal}(\boldsymbol{x}, 1)$$

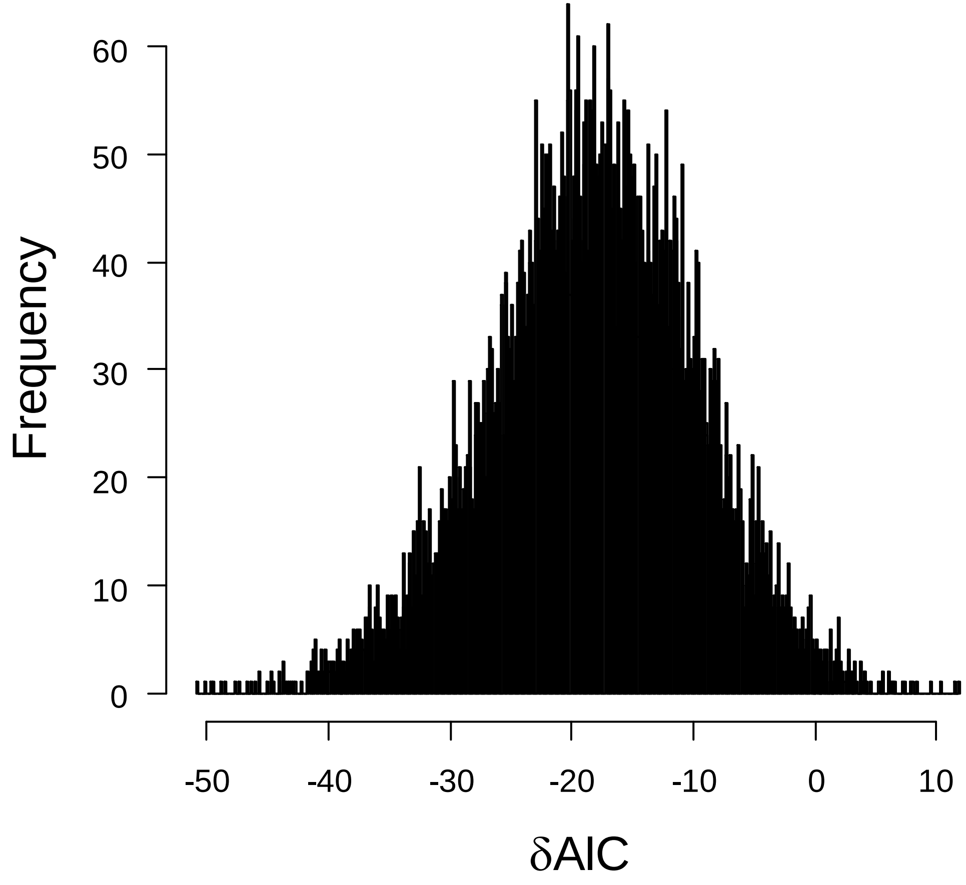
$$\boldsymbol{m} \sim \text{Normal}(\boldsymbol{x}, 1)$$

$$\text{n} = \cancel{25}\ 100$$

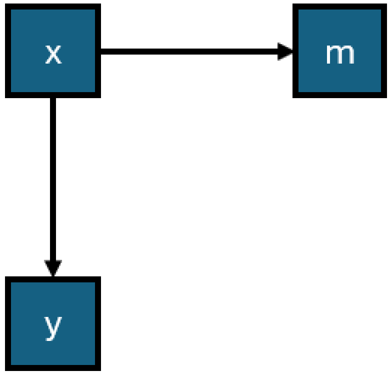
Results (n = 100; 10k sims)



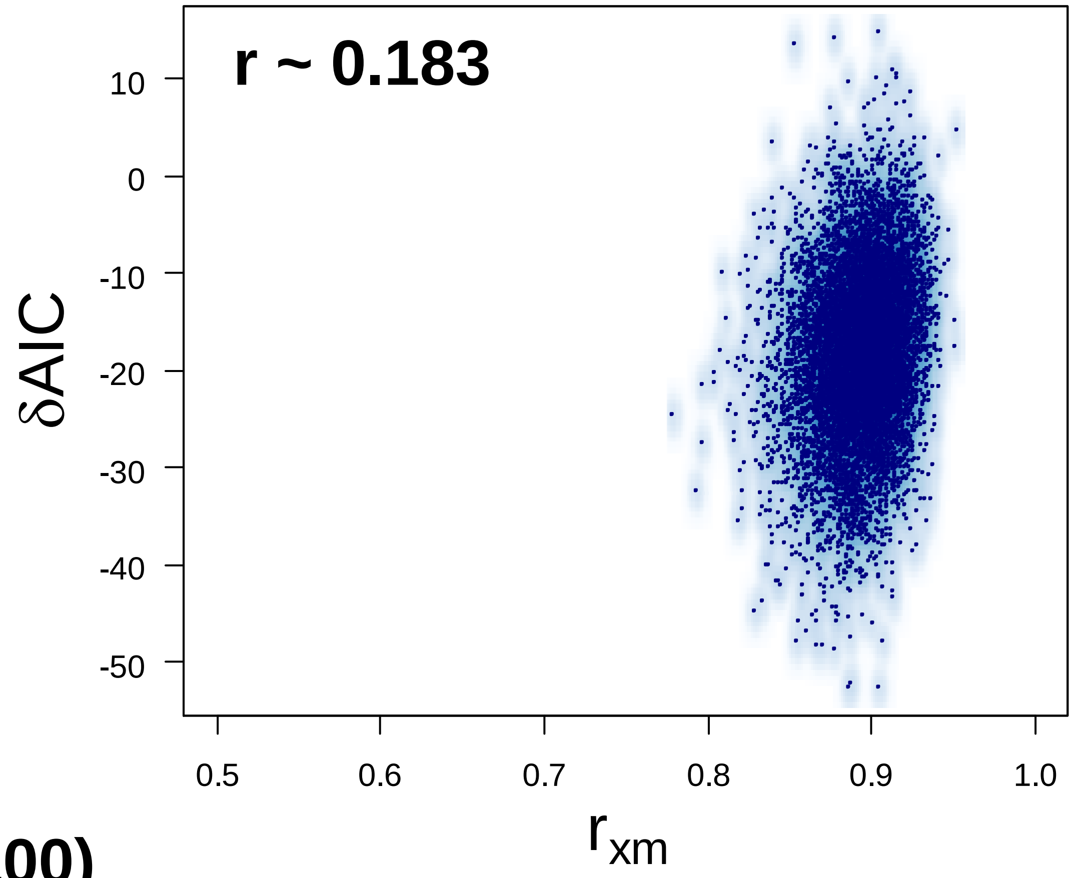
- mean(δAIC) = -18.28
- median(δAIC) = -18.13
- In 1% of simulations, m2 (the spurious model) was superior to m1 (the true model).
- In 2% of simulations, m2 was equivalent to m1 (δAIC < 2).
- m2 had some support (δAIC < 7) in 8% of simulations.



Results (n = 100; 10k sims)



- Increases in correlations among x and m led to difficulty in resolving differences among competing models



It works! (with n = 100)

Ok, deep breath….

There is a clear relationship between the amount of information explained by a covariate and model performance

If a covariate explains information, the parameter estimate will be ‘significant’

If the parameter estimate is significant (at 85% CIs), AIC will improve (decrease). If it doesn’t, then the parameter isn’t doing much…

There is nothing 'wrong' with AIC (or DIC, BIC, WAIC, LOOIC)*, but…

- It is hard to get the 'right' answer and we need large samples sizes (*n*)

Each technique has faced some valid criticism, nothing is 'perfect'

There is nothing 'wrong' with AIC (or DIC, BIC, WAIC, LOOIC)*, but…

- It is hard to get the 'right' answer and we need large samples sizes (*n*)
- In reality, we don't know the 'right' answer (this is why I love data simulation)

Each technique has faced some valid criticism, nothing is 'perfect'

There is nothing 'wrong' with AIC (or DIC, BIC, WAIC, LOOIC)*, but…

- It is hard to get the 'right' answer and we need large samples sizes (*n*)
- In reality, we don't know the 'right' answer (this is why I love data simulation)
- If we're interested in annual estimates, large *n*'s are hard to obtain. Each additional year leads to *n*+1

Each technique has faced some valid criticism, nothing is 'perfect'

There is nothing 'wrong' with AIC (or DIC, BIC, WAIC, LOOIC)*, but…

- It is hard to get the 'right' answer and we need large samples sizes (*n*)
- In reality, we don't know the 'right' answer (this is why I love data simulation)
- If we're interested in annual estimates, large *n*'s are hard to obtain. Each additional year leads to *n*+1
- These are simple univariate regression models! Model selection <u>will not save you</u> from 'bad ideas' in complex systems.

***Each technique has faced some valid criticism, nothing is 'perfect'**

There is nothing 'wrong' with AIC (or DIC, BIC, WAIC, LOOIC)*, but…

- It is hard to get the 'right' answer and we need large samples sizes (*n*)
- In reality, we don't know the 'right' answer (this is why I love data simulation)
- If we're interested in annual estimates, large *n*'s are hard to obtain. Each additional year leads to *n*+1
- These are simple univariate regression models! Model selection <u>will not save you</u> from 'bad ideas' in complex systems.
- AIC/AICc (+ others) are only comparable if the same response variables are used!

***Each technique has faced some valid criticism, nothing is 'perfect'**

Bayesian p-values.

Let’s introduce a bit more complexity ☺

We're going to simulate clutch size as a function of body size

$$n = 200$$

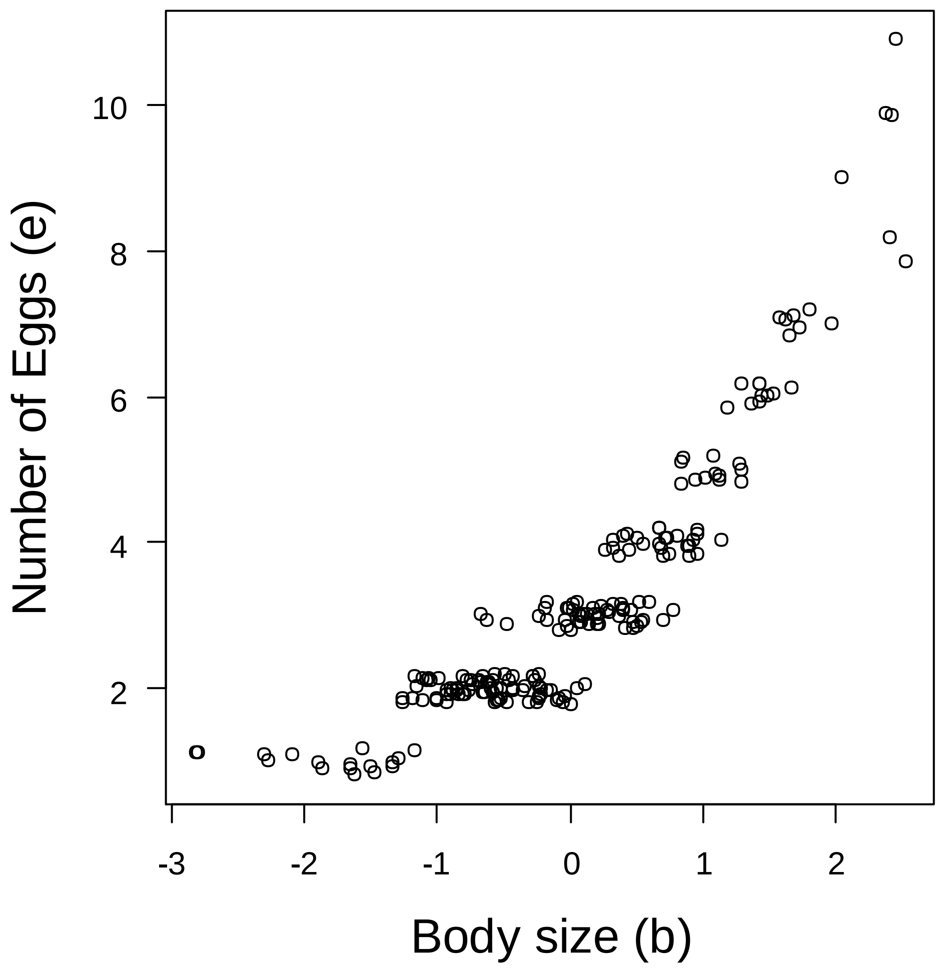
$$\boldsymbol{b} \sim \text{Normal}(0,1)$$

$$\boldsymbol{\psi} = e^{\alpha + \beta \times \boldsymbol{b}}$$

$$\boldsymbol{e} \sim \text{round}(\psi)$$

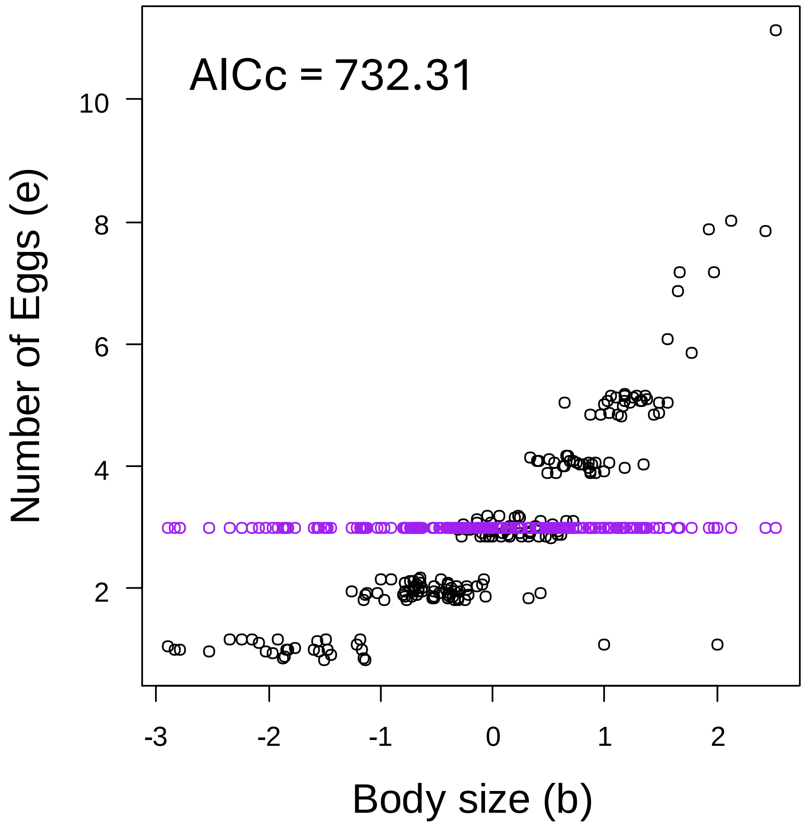
$$\alpha = 1$$

$$\beta = 0.5$$

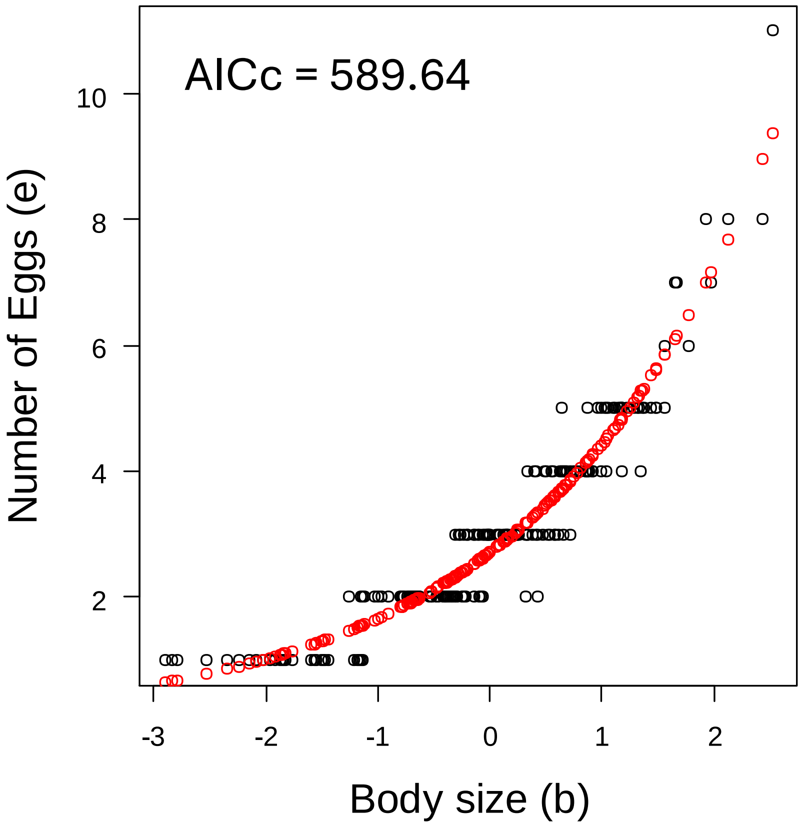


Residuals and AICc…

`glm(y ~ 1, family = 'poisson')`



`glm(y ~ b, family = 'poisson')`



Deviance (D) Information Criterion (DIC; run the same models in JAGS)

Model 0

$$\boldsymbol{\psi} = e^{\alpha}$$

$$\boldsymbol{e} \sim \text{Poisson}(\psi)$$

$$\alpha \sim \text{Normal}(1,1)$$

Model 1

$$\boldsymbol{\psi} = e^{\alpha+\beta\times\boldsymbol{b}}$$

$$\boldsymbol{e} \sim \text{Poisson}(\psi)$$

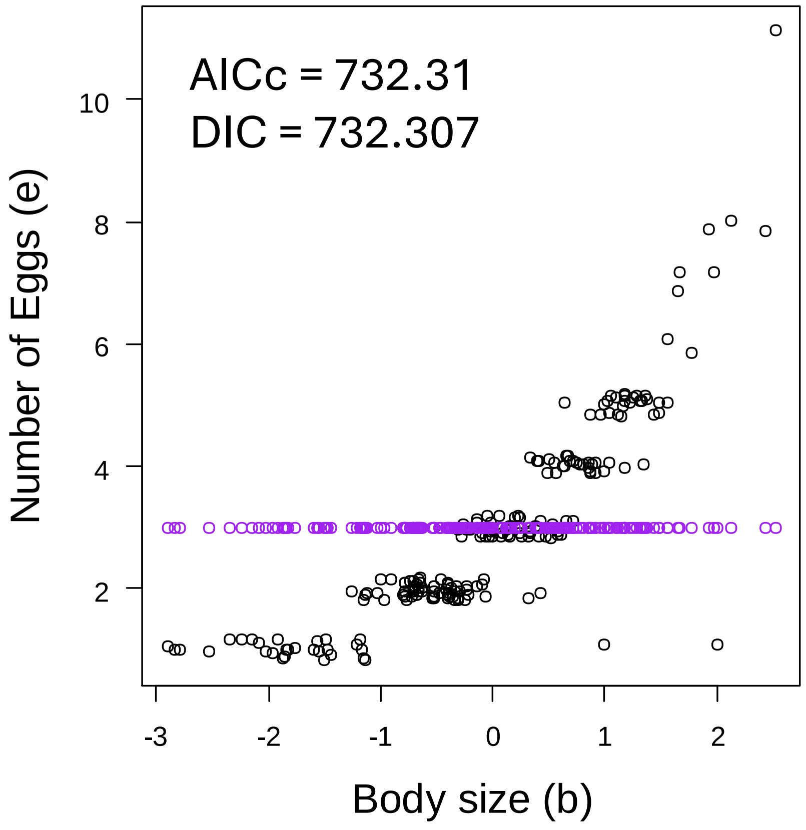
$$\alpha \sim \text{Normal}(1,1)$$

$$\beta \sim \text{Normal}(0,1)$$

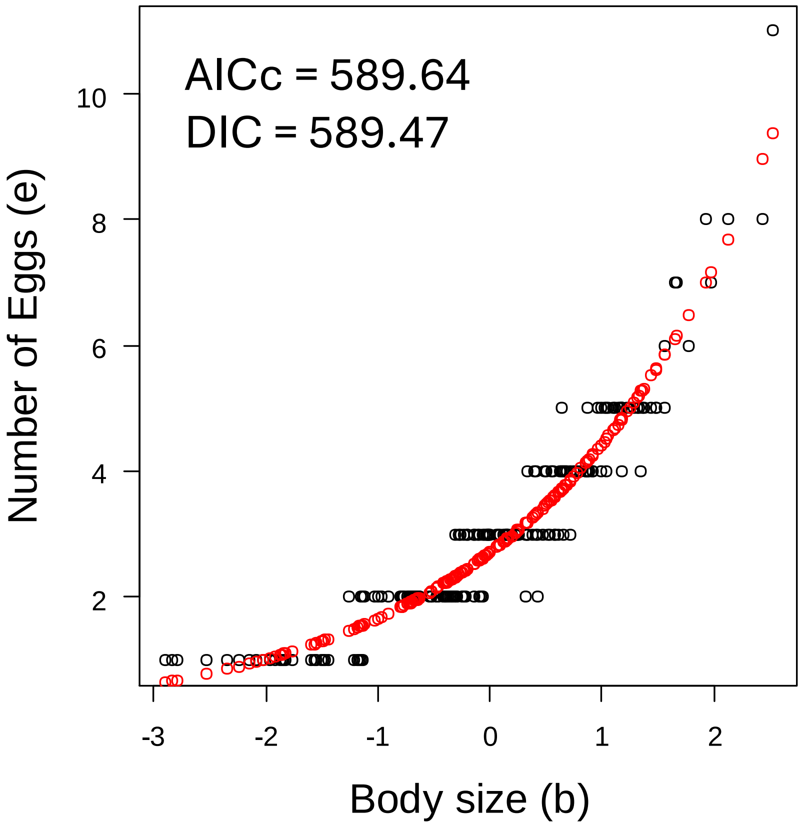
$$\textbf{DIC} = \overline{\boldsymbol{D}} + 2\boldsymbol{pD}$$

$$\boldsymbol{pD} = \frac{\boldsymbol{\sigma_D^2}}{\mathbf{2}}$$

DIC

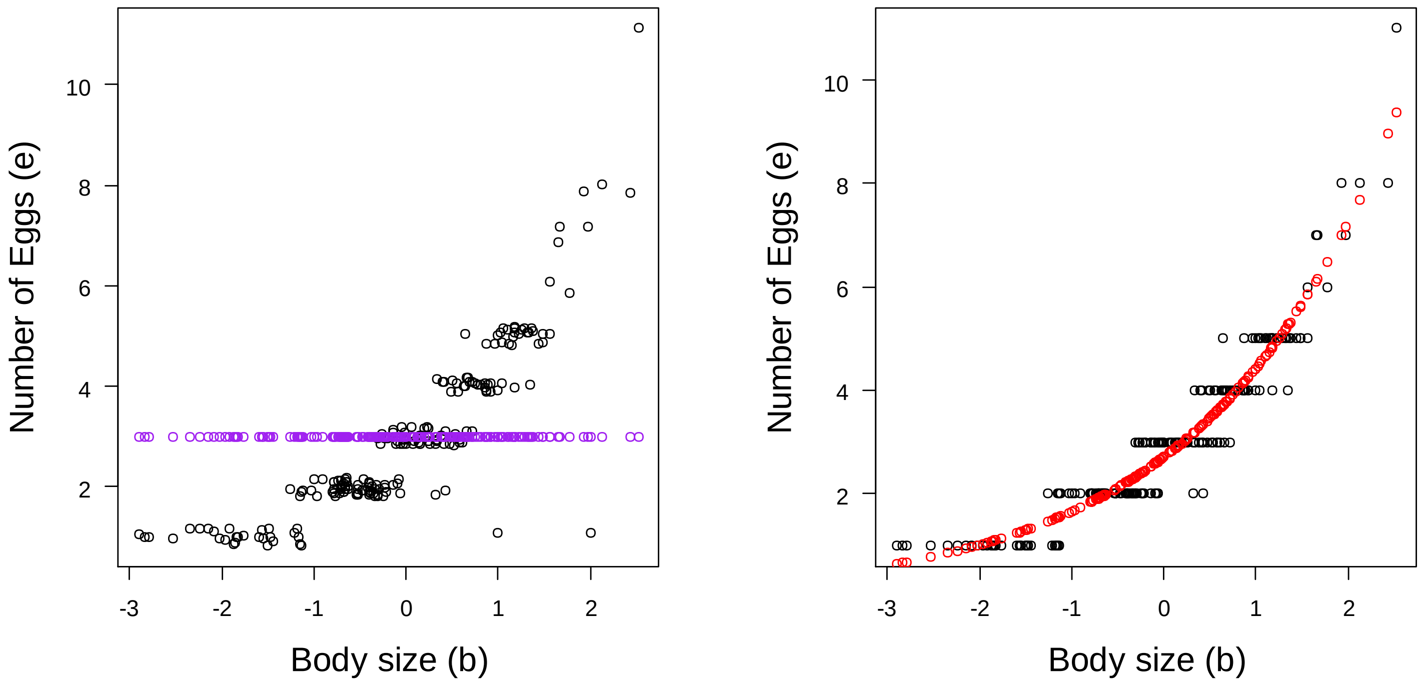


AICc = 732.31
DIC = 732.307
Number of Eggs (e)
Body size (b)



AICc = 589.64
DIC = 589.47
Number of Eggs (e)
Body size (b)

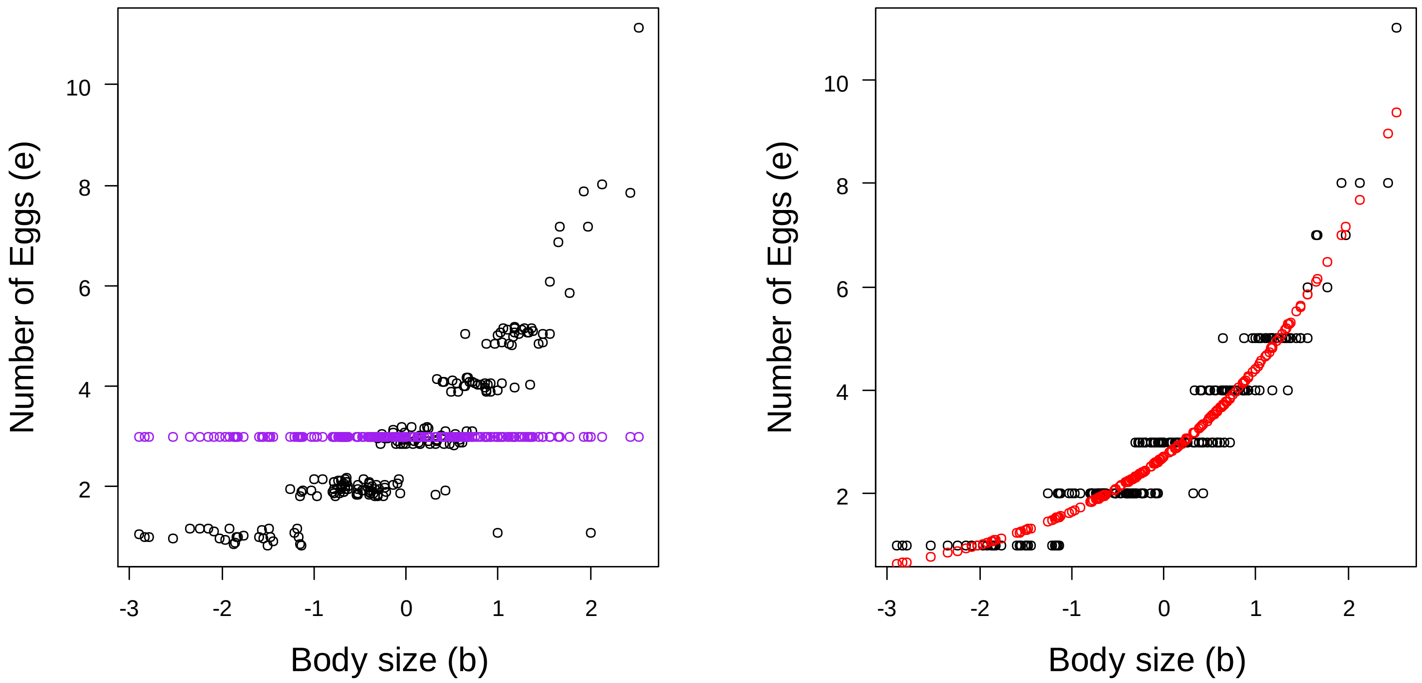
Bayesian p-values (RSS for this Poisson regression)



1. **Calculate (Pearson's) residuals at each iteration**

$$\boldsymbol{\varepsilon} = \frac{\boldsymbol{e} - e^{\beta_0 + \beta_1 \times \boldsymbol{b}}}{\sigma_e}$$

Bayesian p-values (RSS for this Poisson regression)

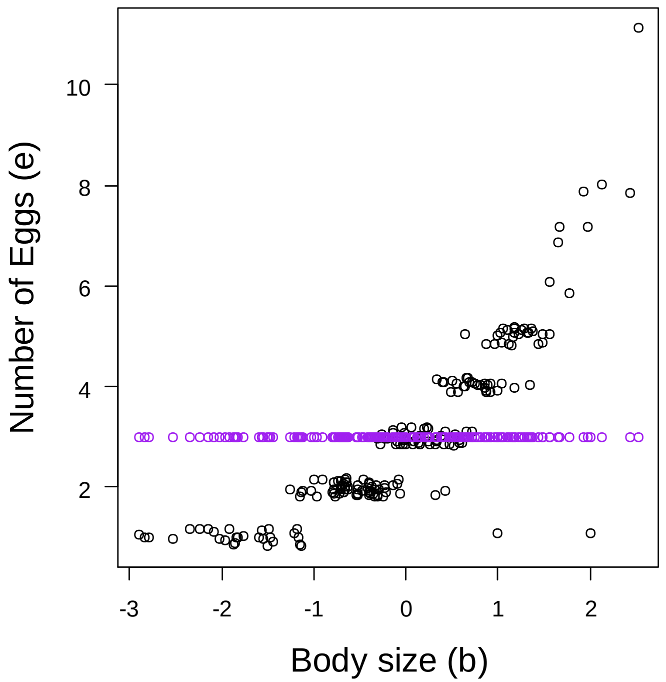


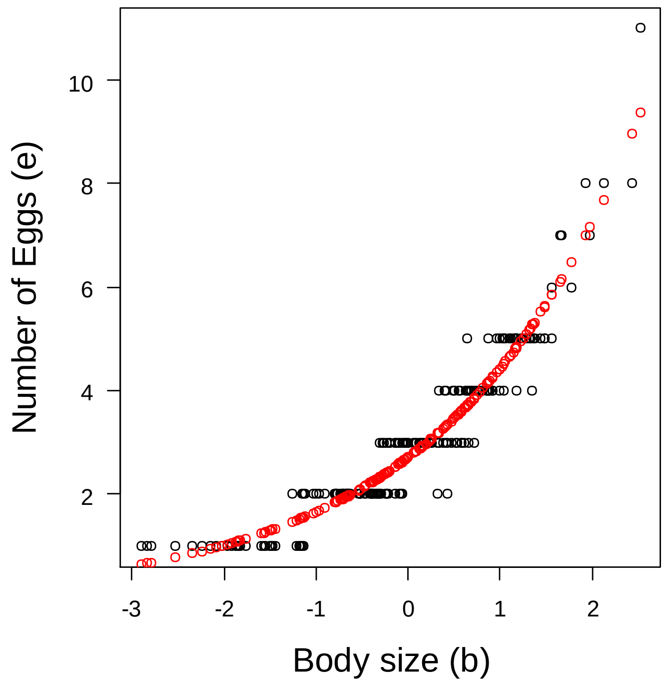
1. **Calculate (Pearson's) residuals at each iteration**
2. **Generate 'new' data**

$$\boldsymbol{e} \sim \mathrm{Poisson}(\boldsymbol{\psi})$$

$$\boldsymbol{e}' \sim \mathrm{Poisson}(\boldsymbol{\psi})$$

Bayesian p-values (RSS for this Poisson regression)

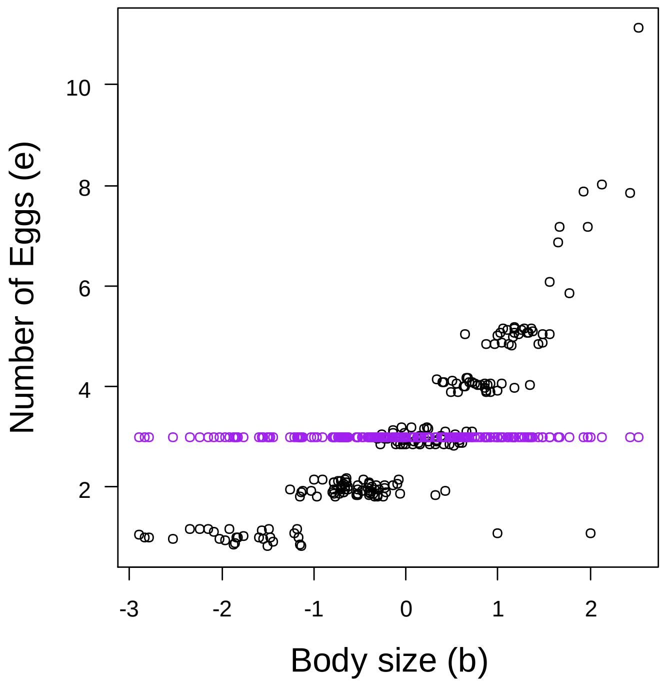


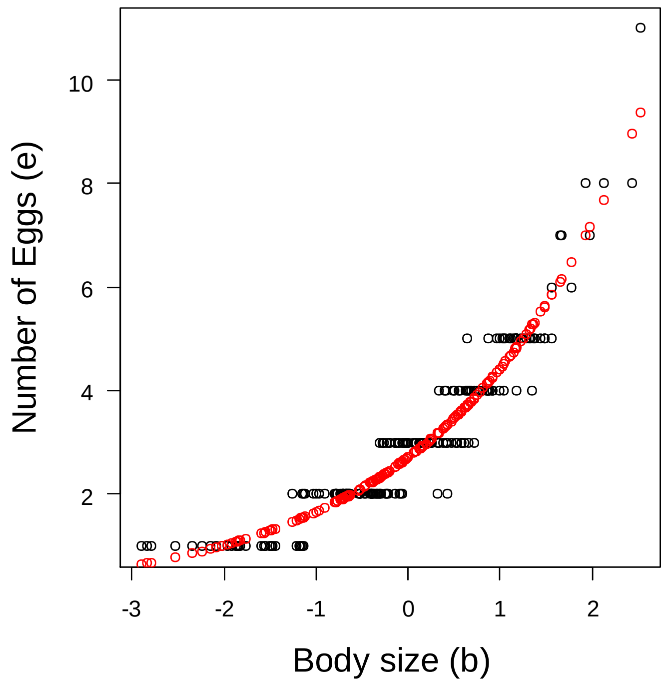


1. **Calculate (Pearson's) residuals at each iteration**
2. **Generate 'new' data**
3. **Calculate (Pearson's) residuals at each iteration for new data**

$$\varepsilon' = \frac{e' - e^{\beta_0 + \beta_1 \times b}}{\sigma_{e'}}$$

Which model will fit the data better?

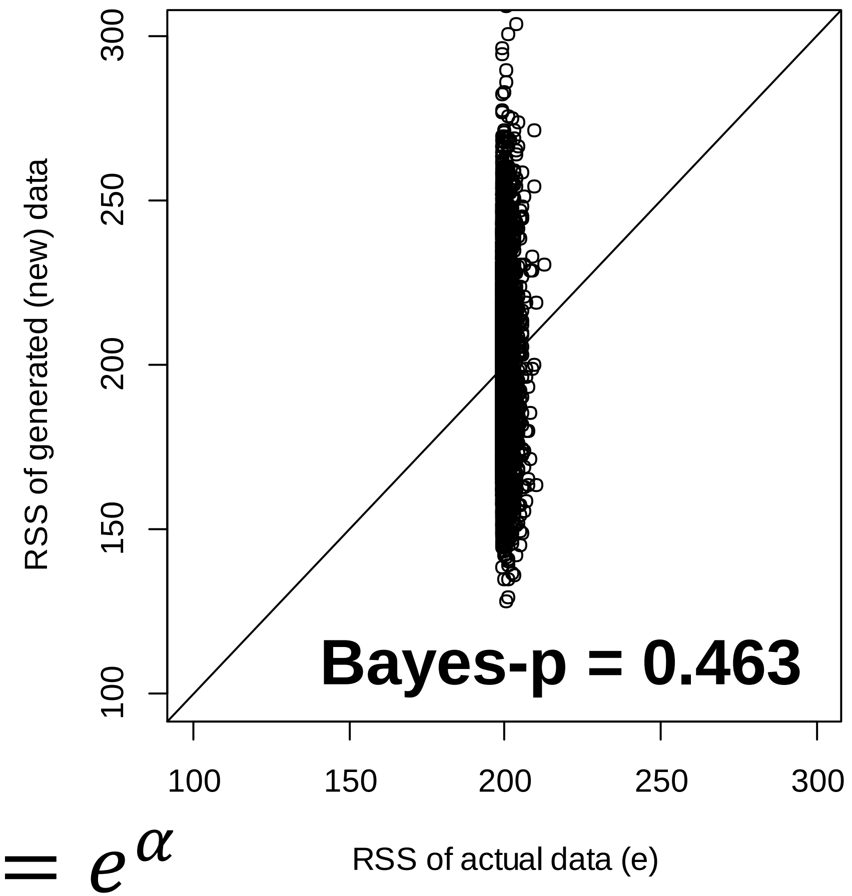


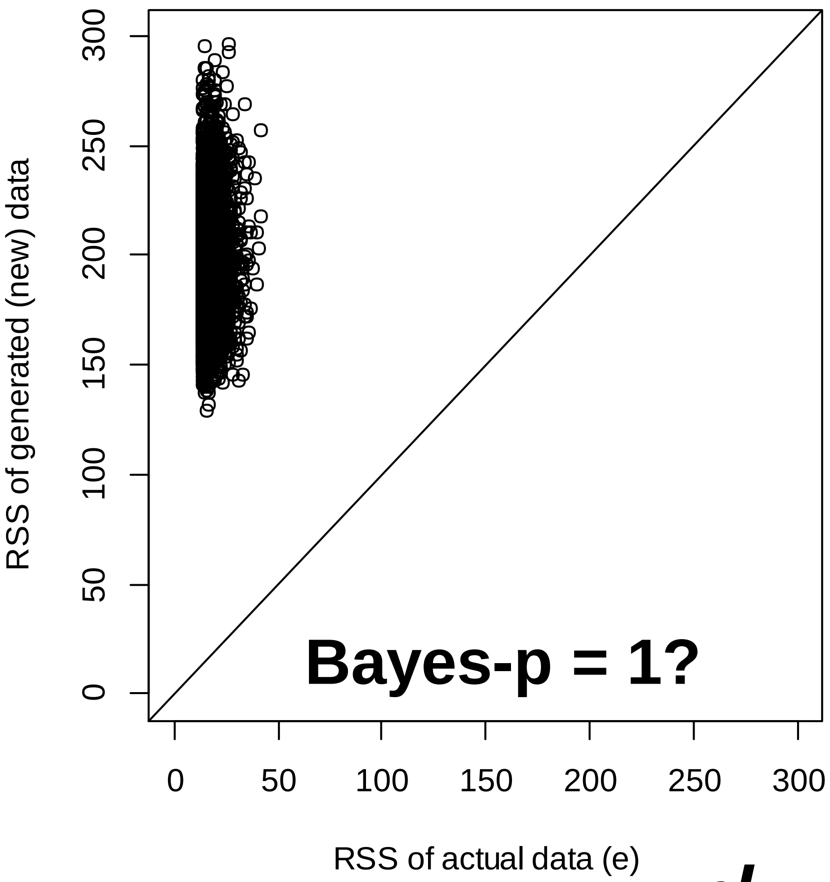


1. **Calculate (Pearson's) residuals at each iteration**
2. **Generate 'new' data**
3. **Calculate (Pearson's) residuals at each iteration for new data**

$$\boldsymbol{\varepsilon}' = \frac{\boldsymbol{e}' - e^{\beta_0 + \beta_1 \times \boldsymbol{b}}}{\sigma_{e'}}$$

Oh for goodness sake!!





$$\boldsymbol{\psi} = e^{\alpha}$$

$$\boldsymbol{e} \sim \mathrm{Poisson}(\psi)$$

$$\boldsymbol{\psi} = e^{\alpha+\beta\times\boldsymbol{b}}$$

$$\boldsymbol{e} \sim \mathrm{Poisson}(\psi)$$