Today
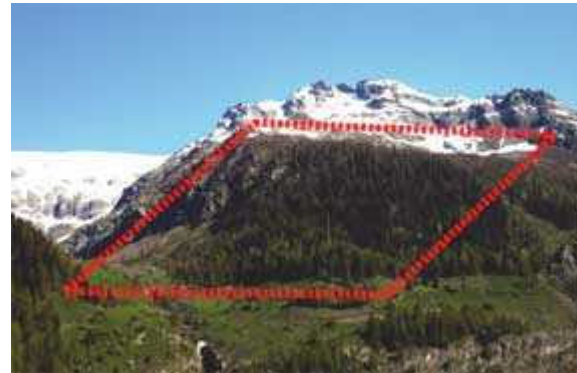
- Why does modeling groups as stochastic matter?

- Binomial GLM

Binomial GLM

- Swiss breeding bird survey (www.vogelwarte.ch)
- Skilled observers, 1 km^2 cells
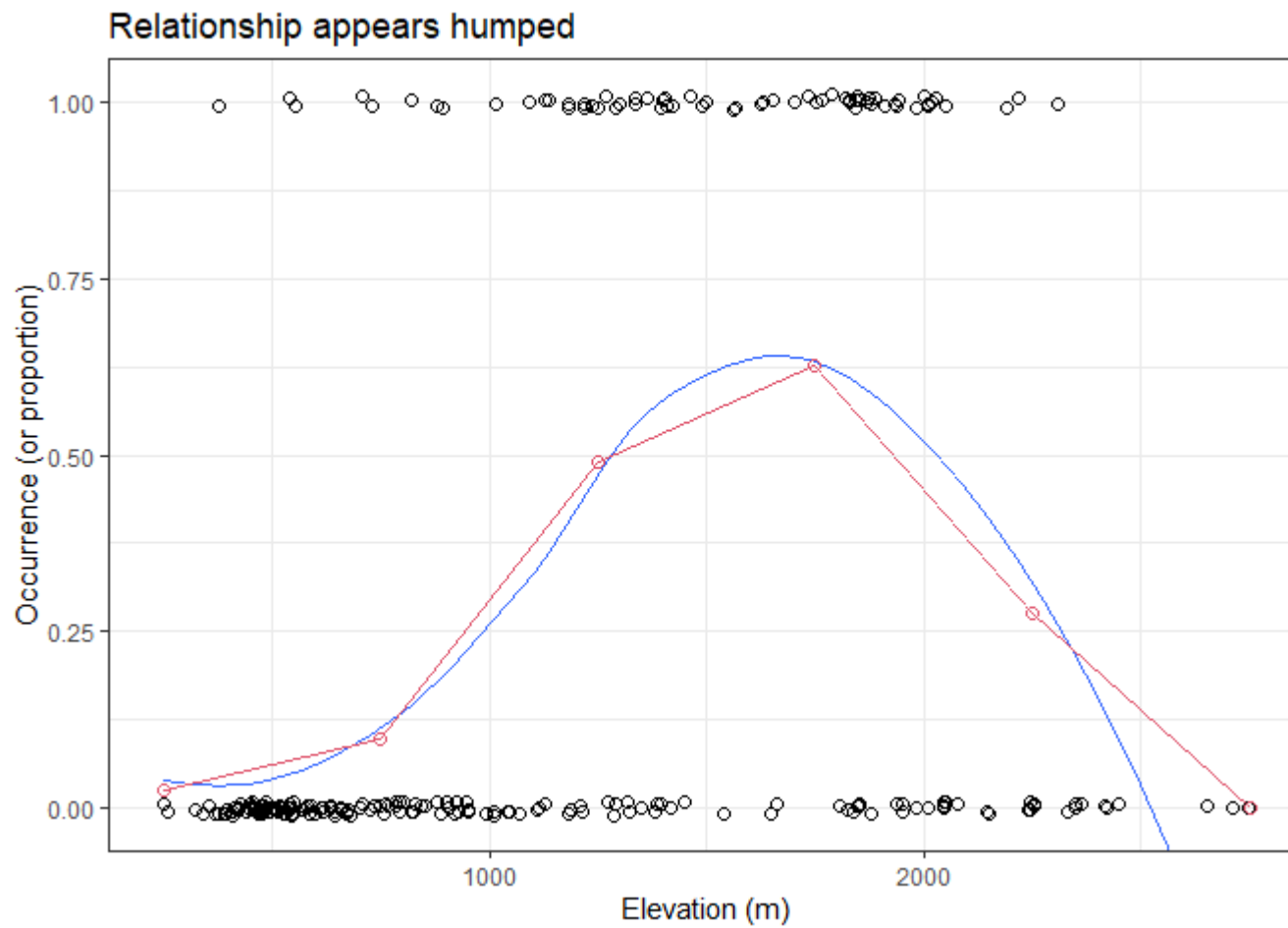

Switzerland; showing
survey locations

A 1 km^2 survey cell

- Willow tit territory presence-absence in relation to altitude

see swissbbs.R

Royle JA, Dorazio RM (2008) Hierarchical Modeling and
Inference in Ecology. Academic Press, Oxford. p 87.

EDA

Relationship appears humped

The model

- Linear predictor; logit link function Biology/pattern

Logit function

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x + \beta_2 x^2$$

Quadratic function allows for hump

Elevation

$$p = \frac{\exp(\text{logit}(p))}{1 + \exp(\text{logit}(p))}$$

Antilogit function: backtransforms to give probability p

The model

- Linear predictor; logit link function Biology/pattern

Logit function

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x + \beta_2 x^2$$

Quadratic function allows for hump

Elevation

$$p = \frac{\exp(\text{logit}(p))}{1 + \exp(\text{logit}(p))}$$

Antilogit function: backtransforms to give probability *p*

- Data model (expected biology + stochasticity)

$$Y \sim \text{Binomial}(p, N = 1)$$

Number of trials = 1

Y is occurrence

Model algorithm

- ## For the probability

```
p_pred_quadratic  <- function(b0, b1, b2, elev){
       lp <- b0 + b1 * elev + b2 * elev^2    #logit p
       prob <- exp(lp) / (1 + exp(lp))       #antilogit
       return(prob)
}
```

- ## For data generating process

```
rbinom(n, size=1, prob=p_pred_quadratic(b0,b1,b2,elev))
```

Training algorithm

- Maximum likelihood

Likelihood for a single data point

$$L(p) = P(Y = y|p) = p^y(1-p)^{1-y}$$

1 for present
0 for absent

Bernoulli distribution

y = 1:

$$L(p) = P(Y = 1|p) = p^1(1-p)^{1-1} = p$$

y = 0:

$$L(p) = P(Y = 0|p) = p^0(1-p)^{1-0} = 1-p$$

Training algorithm (general way)

```
1  p_pred_quadratic  <- function(b0,b1,b2,elev){
        lp <- b0 + b1 * elev + b2 * elev^2    #logit p
        prob <- exp(lp)/(1+exp(lp))           #antilogit
        return(prob)
   }
```

β_2 is orders of magnitude smaller than other parameters. Rescale to get optim to behave better.

```
2  quadratic_nll <- function(p,occ,elev){

        b2 <- p[3] * 1e-06   #Rescale
        ppred <- p_pred_quadratic(b0=p[1],b1=p[2],b2,elev)
        nll <- -sum(dbinom(occ,size=1,prob=ppred,log=TRUE))
        return(nll)
   }
```

β_2 = -4e-06

```
3  par <- c(-5,0.02,-4) #Starting values
   optim(par,quadratic_nll,occ=occ,elev=elev)
```

Fitted model compared to binned proportions
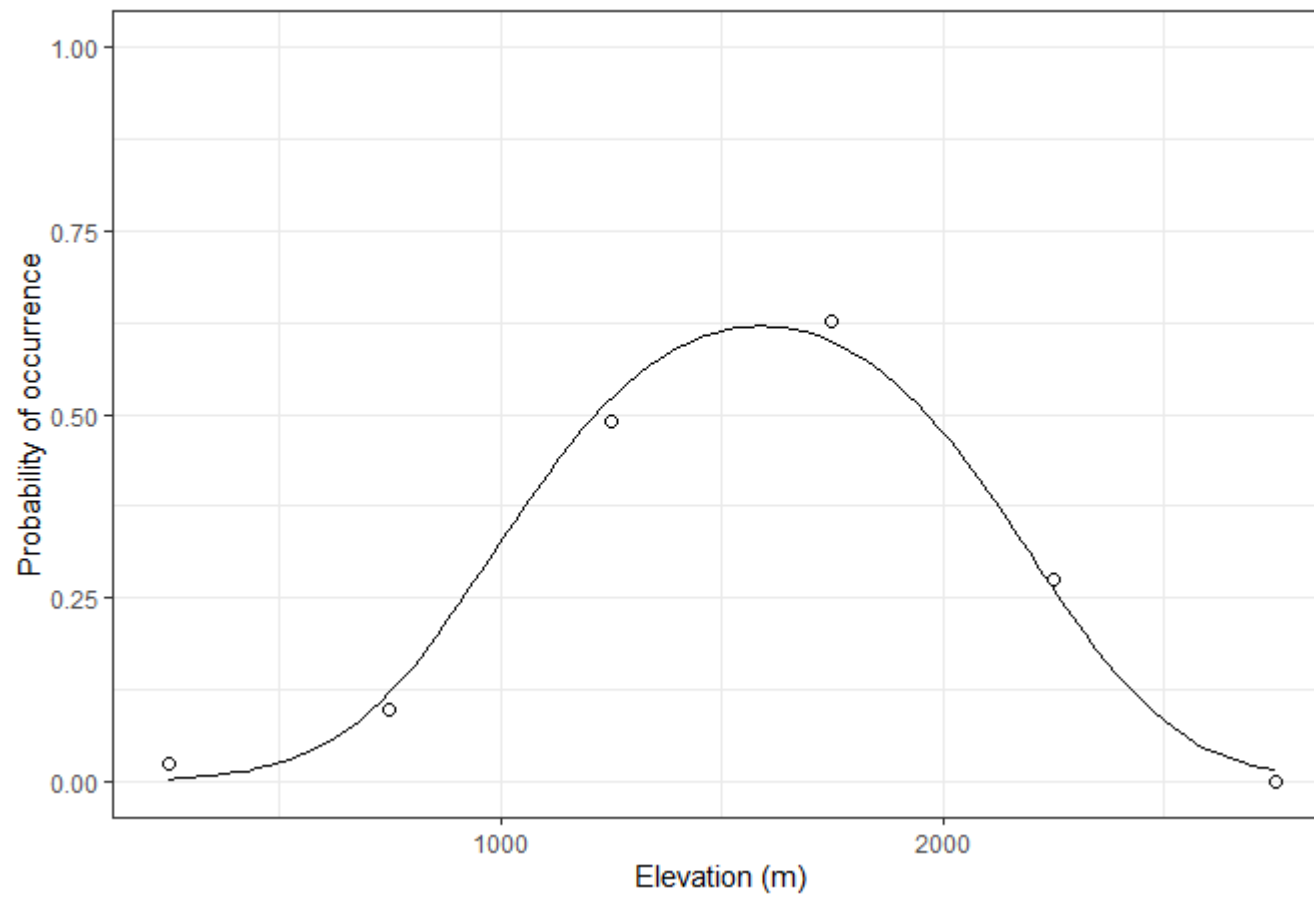
Inference algorithm

Test an ecological hypothesis:
Species' distribution is humped

- H_1: Hump

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x + \beta_2 x^2$$
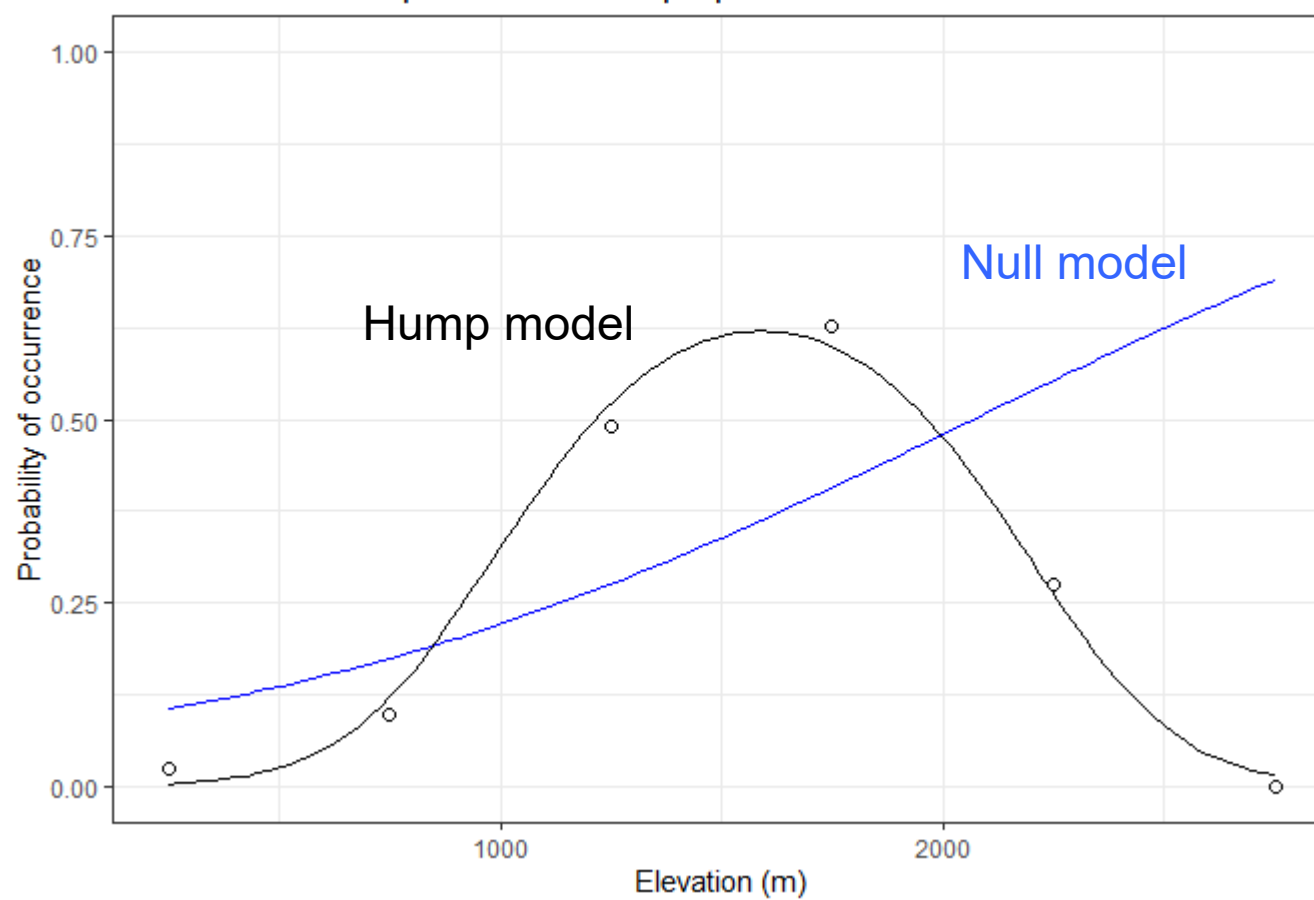
- H_0: No hump

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

Nested models since we can set β_2 to zero and get H_0.

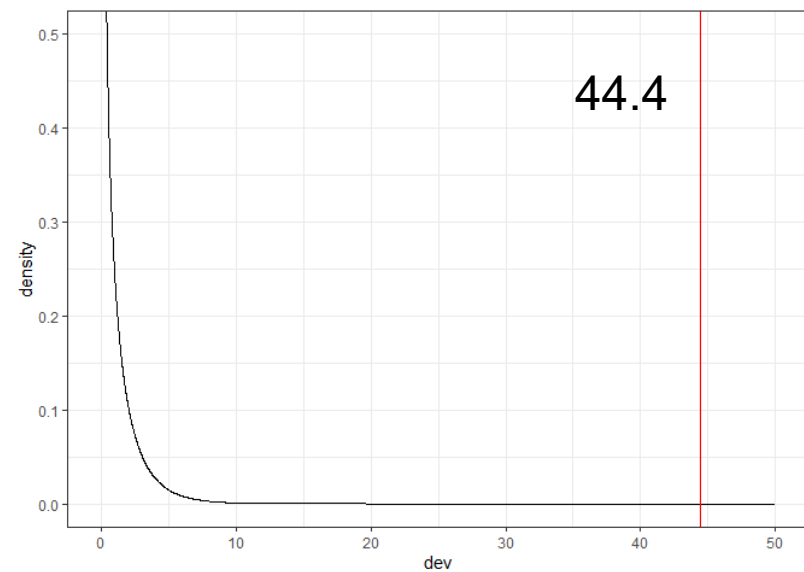Fitted models compared to binned proportions

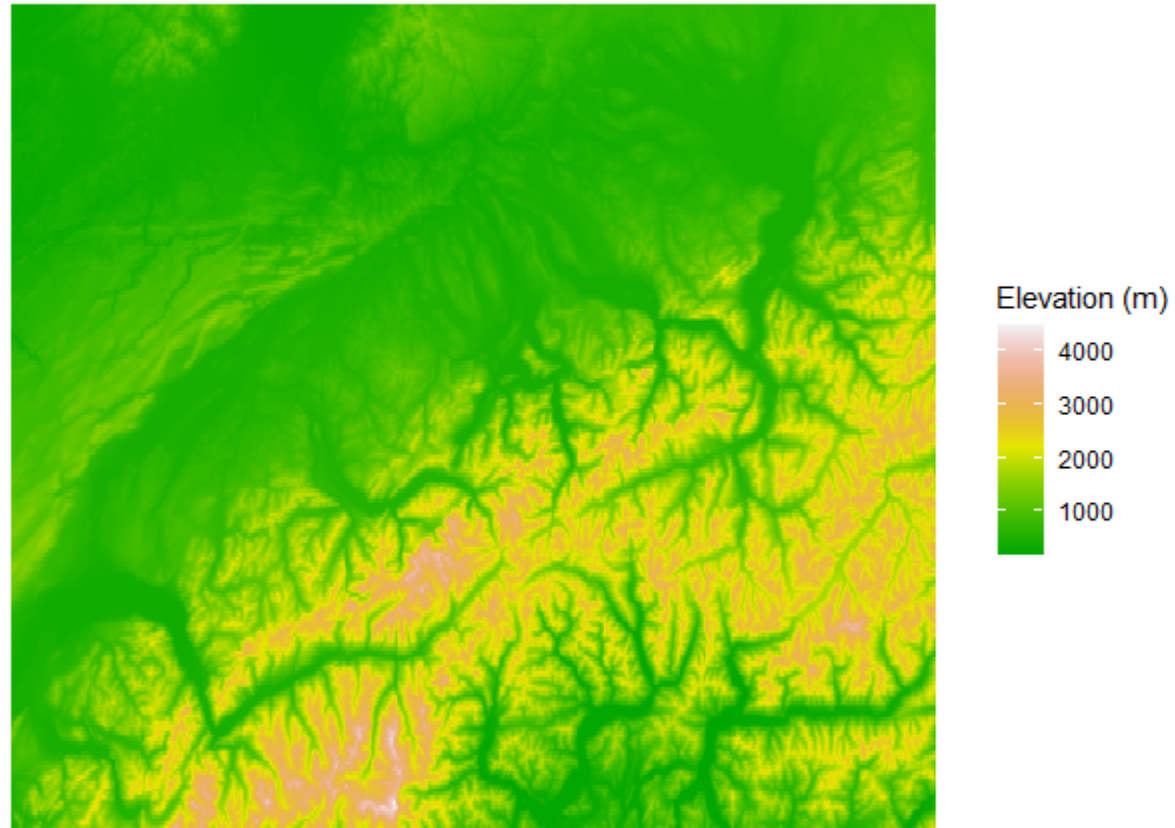# Inference algorithm

- NHST: likelihood ratio test

Test statistic

$$\text{deviance} = 2 \, ln\left(\frac{L_1}{L_0}\right)$$
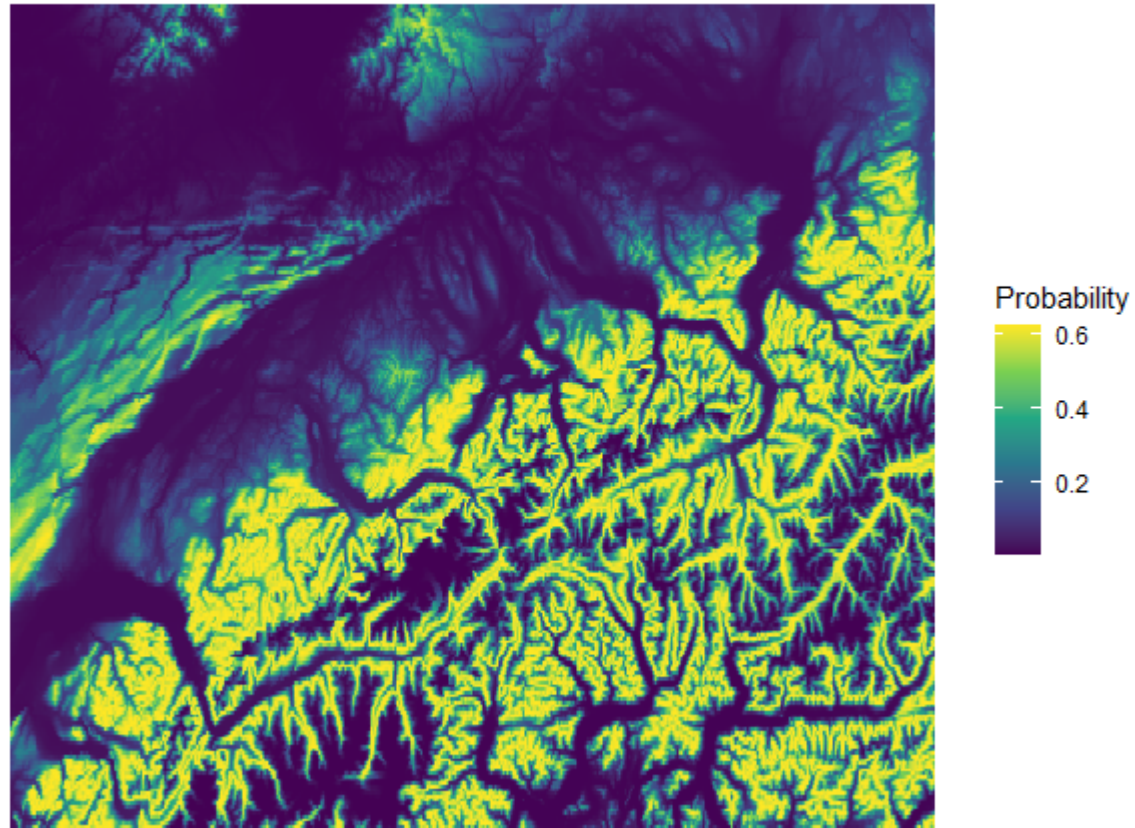
P = 2.7e-11

Sampling distribution (chi-squared)

44.4

Switzerland: DEM

Probability of observing a willow tit in year 1

R code

- See
- swissbbs.R
- swissbbs.md
- also for the easy way:
  - frequentist: glm()
  - Bayesian: stan_glm()