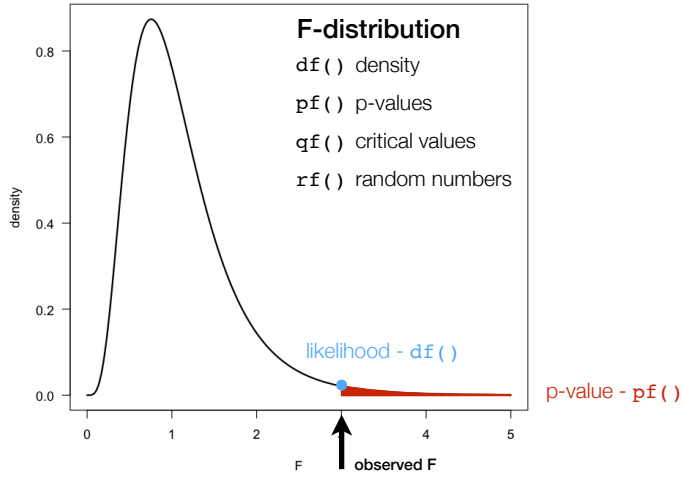
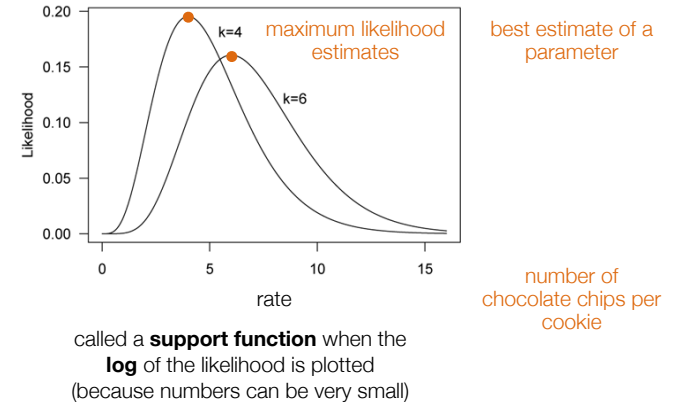


likelihood vs. p-values



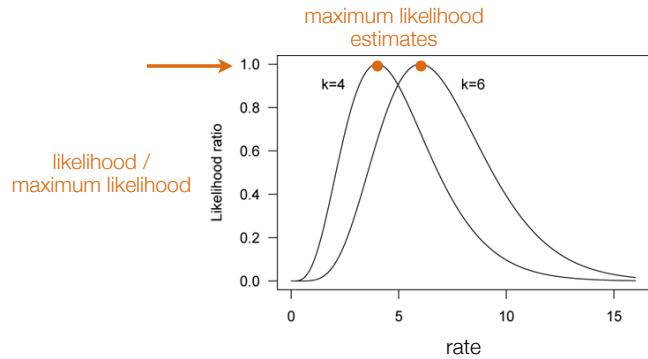
Maximum Likelihood Estimate (MLE)

(at the maximum value on the likelihood curve)



Likelihood Ratio

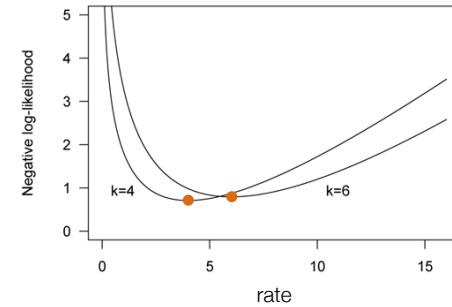
(scaling likelihood curves by the MLE)



Negative Log-Likelihood

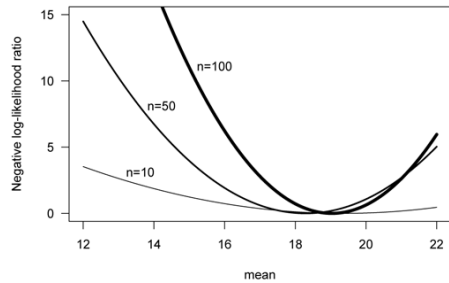
(MLE is therefore at the minimum)

because we generally minimize functions in statistics



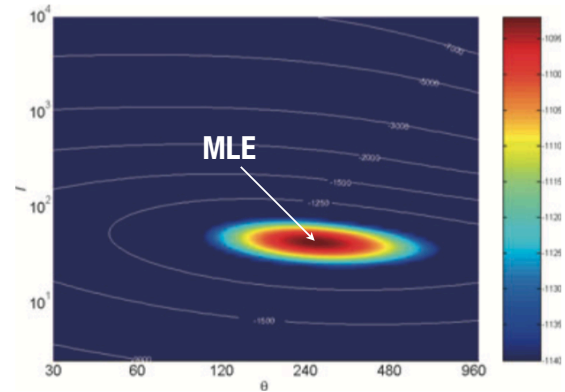
Shape of likelihood curves with sample size

(larger sample sizes decrease uncertainty in the MLE)



can be used to create confidence intervals

Two parameters have a likelihood surface



Model Selection 1

(for nested models)

Likelihood Ratio Test: Simple (s) vs. Complex (c) model

$$R = 2(\overset{\text{log-likelihoods}}{LL_c} - \overset{\text{likelihoods}}{LL_s}) = -2\log_e \left(\frac{L_s}{L_c} \right)$$

p-value from chi-squared distribution

Model Selection 2

(works with nested & non-nested models)

Akaike Information Criterion (AIC)

For models (i) differing in the number of parameters (K)

$$AIC_i = -2L_i + 2K_i$$

AIC for model i log-likelihood for model i parameters in model i

best model has the **lowest** (most negative) AIC

Model Selection 3

(works with nested & non-nested models)

Bayesian Information Criterion (AIC)

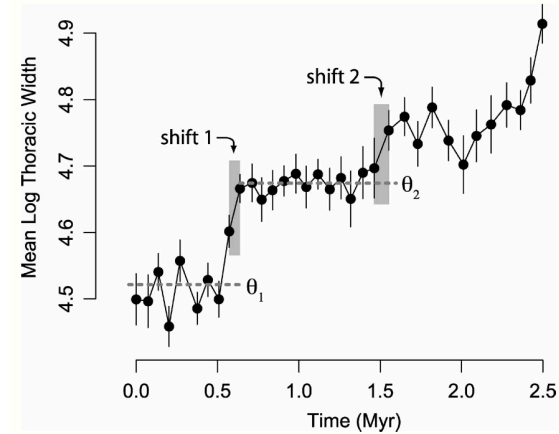
For models (i) differing in the number of parameters (K)

$$BIC_i = -2L_i + K_i(\log(n))$$

BIC for model i log-likelihood for model i parameters in model i sample size

best model has the lowest (most negative) BIC

Example: evolutionary history



- random walk?
- directional?
- stasis?
- one punctuation?
- two punctuations?
- one punctuation, then random walk?

Gene Hunt (2008, Paleobiology 34:360)

TABLE 2. Model fits to *Pseudocubus vema* thoracic width data. All punctuations are of the unsampled variety (see text); models 4 and 5 posit one and two punctuations, respectively. The punctuated models fit here assume the separate evolutionary variances for each interval of stasis. Model 6 postulates a single punctuation, and then a later transition to an unbiased random walk. Parameter estimates for the best-supported model (6): $\theta_1 = 4.52 \log \mu\text{m}$, $\theta_2 = 4.67 \log \mu\text{m}$, $\omega_1 = 0.0017 \log \mu\text{m}^2$, $\omega_2 = 0 \log \mu\text{m}^2$, $\sigma_{\text{step}}^2 = 0.018 \log \mu\text{m}/\text{Myr}$; shift points after the ninth and 22nd samples. Akaike weights for models receiving at least moderate support ($w > 0.2$) are in bold. Abbreviations as for Table 1.

No.	Model	No. of segments	K	log(L)	AIC _c	w
1	Unbiased random walk	1	1	53.71	-105.29	0.000
2	Directional evolution	1	2	54.81	-105.21	0.000
3	Stasis	1	2	27.56	-50.73	0.000
4	Punctuation-1	2	5	47.67	-83.11	0.000
5	Punctuation-2	3	8	69.87	-117.74	0.040
6	Punctuation-1, then unbiased random walk	3	7	71.28	-124.08	0.960

largest model complexity / largest number of parameters / best model (ignoring the number of model parameters)

smallest model complexity / smallest number of parameters / best model (penalizing for the number of model parameters)

Hunt (2008, Paleobiology 34:360)

TABLE 3. Model fits to putative punctuated anagenesis of test shape in the *Globorotalia tumida* lineage. "Unbiased random walk-2" and "Unbiased random walk-3" have two and three segments, respectively, each with separate step variance parameter values. The punctuation models (6, 7) differ as to whether the transition interval exhibits inherently directional change (7) or not (6); both assume equal evolutionary variances before and after punctuation event. Parameter estimates for the best-supported model (6): $\theta_1 = -0.057$ units, $\theta_2 = 0.049$ units, $\omega = 0.00013$ units², $\sigma_{\text{step}}^2 = 0.013$ units²/Myr, shift points after the 17th and 65th samples (see Fig. 4). Akaike weights for models receiving at least moderate support ($w > 0.2$) are in bold. Abbreviations as for Table 1.

No.	Model	No. of segments	K	log(L)	AIC _c	w
1	Unbiased random walk	1	1	232.18	-462.32	0.000
2	Directional evolution	1	2	232.40	-460.67	0.000
3	Stasis	1	2	167.51	-331.02	0.000
4	Unbiased random walk-2	2	3	238.57	-470.86	0.000
5	Unbiased random walk-3	3	5	243.12	-475.55	0.002
6	Sampled punctuation (unbiased random walk transition)	3	6	250.37	-487.76	0.763
7	Sampled punctuation (directional evolution transition)	3	7	250.50	-485.70	0.263

largest model complexity / largest number of parameters / best model (ignoring the number of model parameters)

smallest model complexity / smallest number of parameters / best model (penalizing for the number of model parameters)

Hunt (2008, Paleobiology 34:360)

Bayes Rule

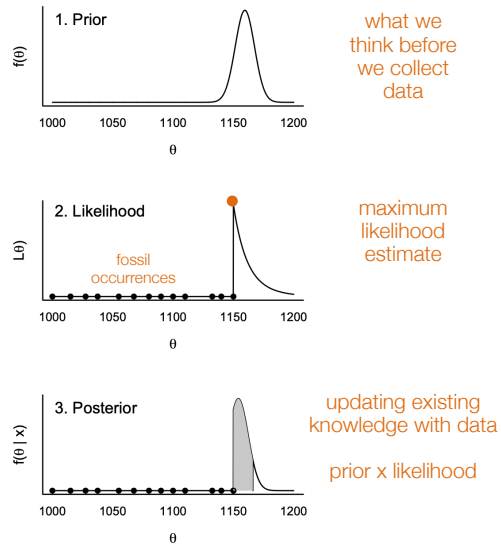
$$P(H|D) = \frac{\overset{\text{Likelihood}}{\underset{\text{probability of the data, given the hypothesis}}{P(D|H)}} \underset{\text{probability of the hypothesis}}{P(H)}}{\underset{\text{probability of the data}}{P(D)}}$$

probability of the hypothesis, given the data

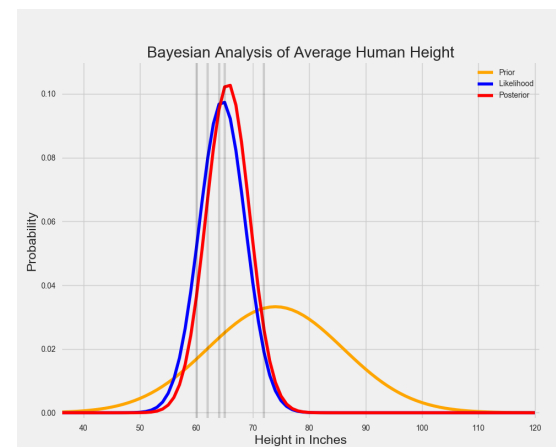
Bayes Rule

$$P(H|D) \propto \overset{\text{posterior probability}}{P(H|D)} \overset{\text{likelihood}}{P(D|H)} \overset{\text{prior probability}}{P(H)}$$

Estimating the time of extinction (θ)

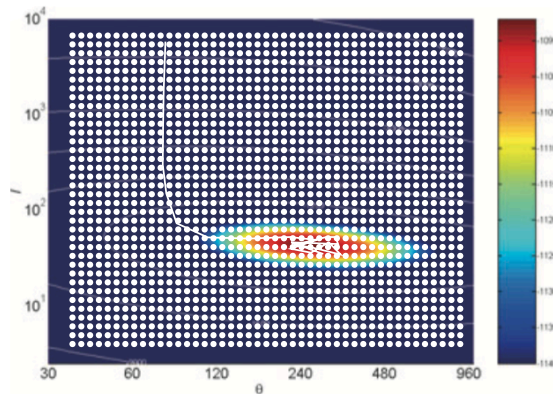


Bayesian problems are simple when the distributions are simple



Strategy 1: calculate likelihood for every combination of parameters

Strategy 2: random walk, looking for higher likelihoods



Random walk would be repeated many times to avoid local maxima/minima

