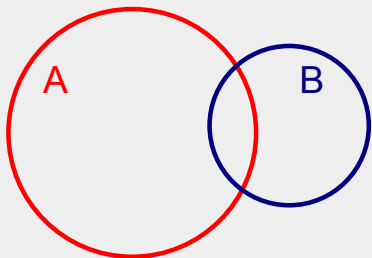
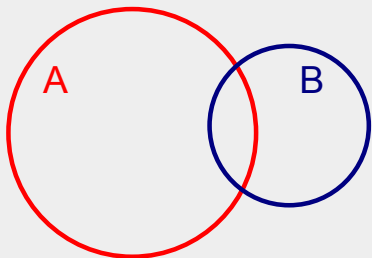


Условная вероятность (conditional probability)



$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Leftrightarrow \quad P(A \cap B) = P(A|B) \cdot P(B)$$

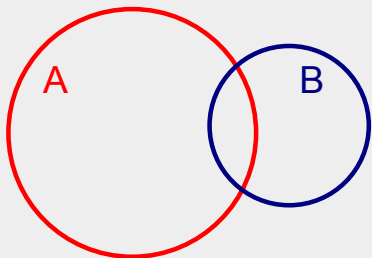
Условная вероятность (conditional probability)



$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Leftrightarrow \quad P(A \cap B) = P(A|B) \cdot P(B)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \Leftrightarrow \quad P(A \cap B) = P(B|A) \cdot P(A)$$

Условная вероятность (conditional probability)



$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Leftrightarrow \quad P(A \cap B) = P(A|B) \cdot P(B)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \Leftrightarrow \quad P(A \cap B) = P(B|A) \cdot P(A)$$

$$P(A|B) \cdot P(B) = P(A) \cdot P(B|A)$$

Теорема Байеса (Bayes' theorem)

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Теорема Байеса (Bayes' theorem)

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

Теорема Байеса (Bayes' theorem)

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

$P(A|B)$ — апостериорная (posterior) вероятность;

$P(A)$ — априорная (prior) вероятность;

$P(B|A)$ — вероятность В при условии А;

$P(B)$ — вероятность события В.

Теорема Байеса (Bayes' theorem)

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

$P(A|B)$ — апостериорная (posterior) вероятность;

$P(A)$ — априорная (prior) вероятность;

$P(B|A)$ — вероятность В при условии А;

$P(B)$ — вероятность события В.

$$P(B) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})$$

Теорема Байеса и филогенетика

$$P(T, \beta, k|X) = \frac{P(T, \beta, k) \cdot P(X|T, \beta, k)}{P(X)}$$

T — топология;

β — длина ветвей;

k — параметры эволюционной модели.

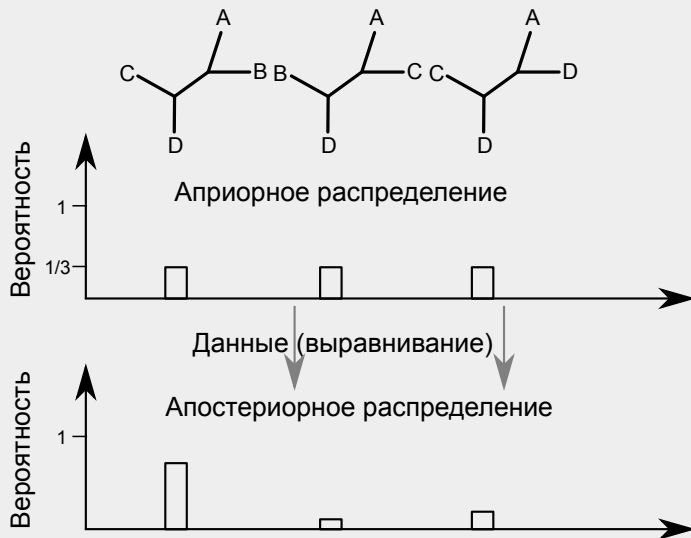
$P(T, \beta, k|X)$ — апостериорная вероятность гипотезы;

$P(T, \beta, k)$ — априорная вероятность гипотезы;

$P(X|T, \beta, k)$ — правдоподобие данных;

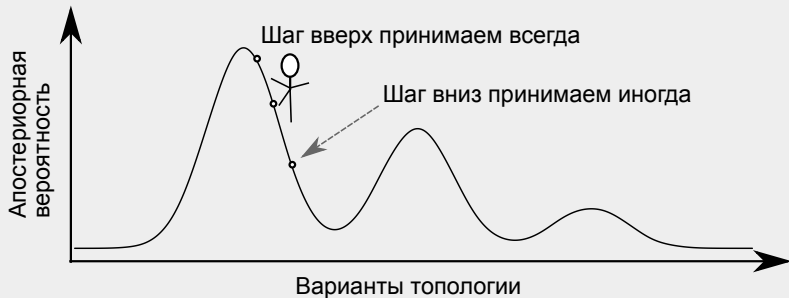
$P(X)$ — нормализующая константа.

Теорема Байеса и филогенетика



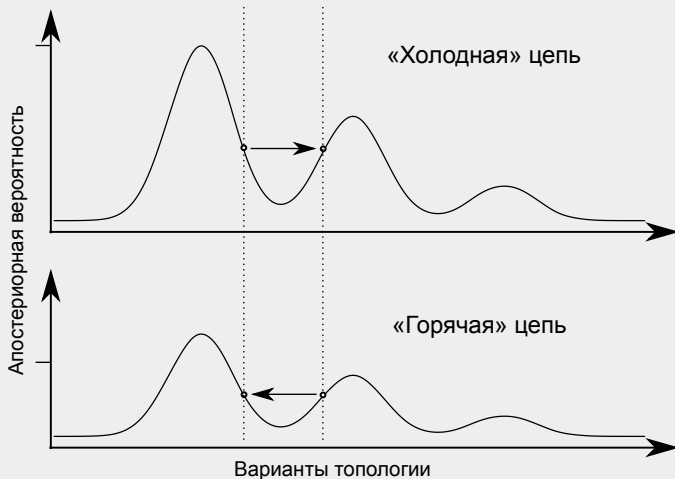
По Lemey, P, Salemi, M. and Vandamme, A. M. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. Cambridge University Press, 2009.

MCMC (Markov chain Monte Carlo, алгоритм Монте-Карло по схеме марковской цепи)



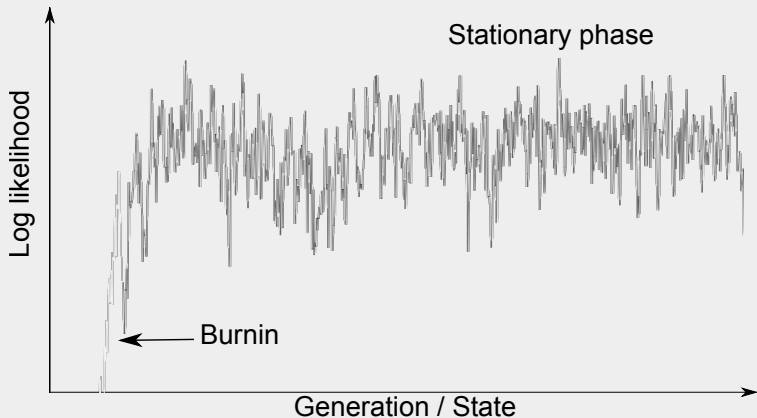
По Lemey, P, Salemi, M. and Vandamme, A. M. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. Cambridge University Press, 2009.

MCMCMC (Metropolis coupled MCMC)



По Lemey, P, Salemi, M. and Vandamme, A. M. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. Cambridge University Press, 2009.

Burn-in and convergence (сходимость)



www.molrev.org/resources/activities/beast_activity/viruses

см. также Lemey, P, Salemi, M. and Vandamme, A. M. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. Cambridge University Press, 2009.

Априорные распределения

$$P(T, \beta, k|X) = \frac{P(T, \beta, k) \cdot P(X|T, \beta, k)}{P(X)}$$

T — топология;

β — длина ветвей;

k — параметры эволюционной модели.

Априорные распределения

$$P(T, \beta, k|X) = \frac{P(T, \beta, k) \cdot P(X|T, \beta, k)}{P(X)}$$

T — топология;

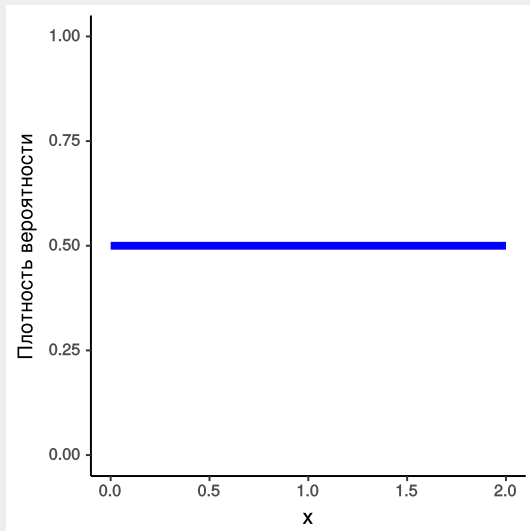
β — длина ветвей;

k — параметры эволюционной модели.

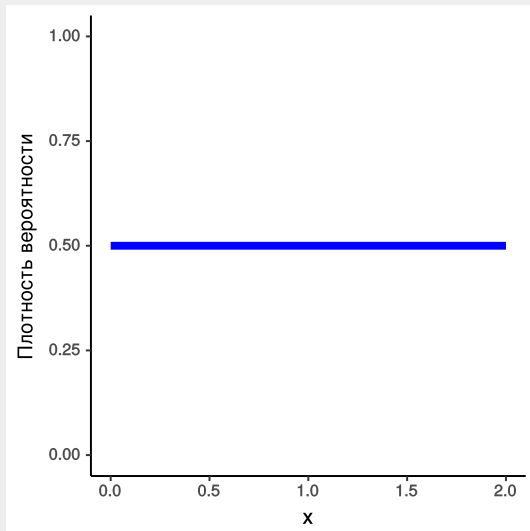
Наиболее часто используемые распределения:

- ▶ Равномерное
- ▶ Экспоненциальное
- ▶ Гамма-распределение
- ▶ Бета-распределение и распределение Дирихле

Равномерное распределение

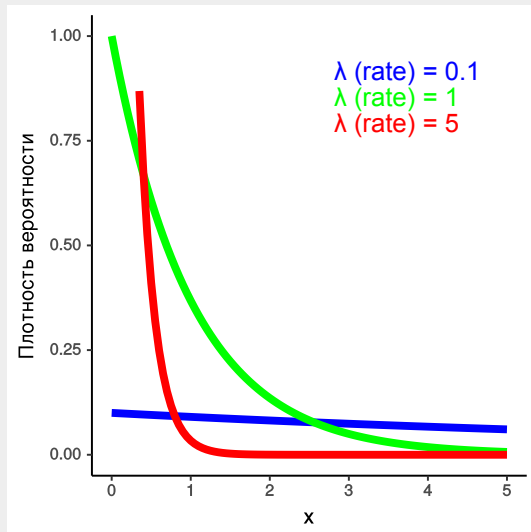


Равномерное распределение

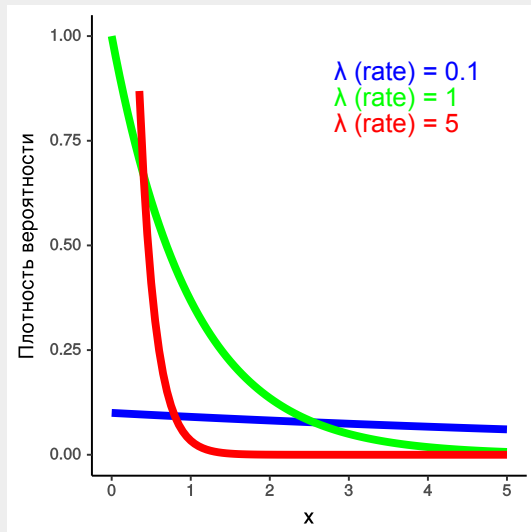


Топология

Экспоненциальное распределение

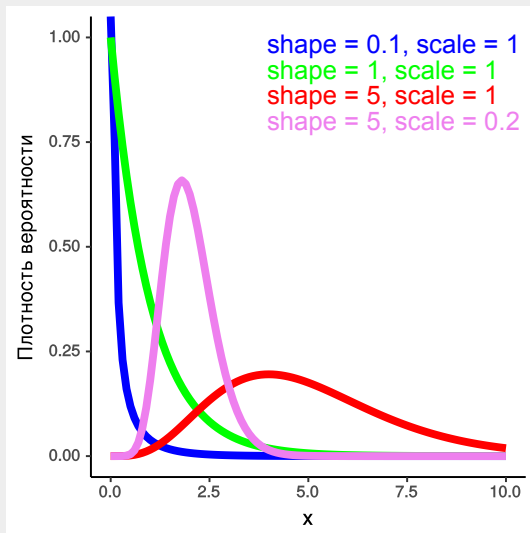


Экспоненциальное распределение

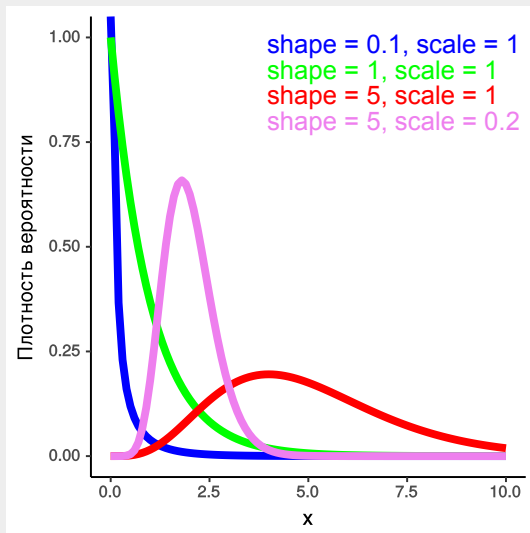


Длина ветвей

Гамма-распределение

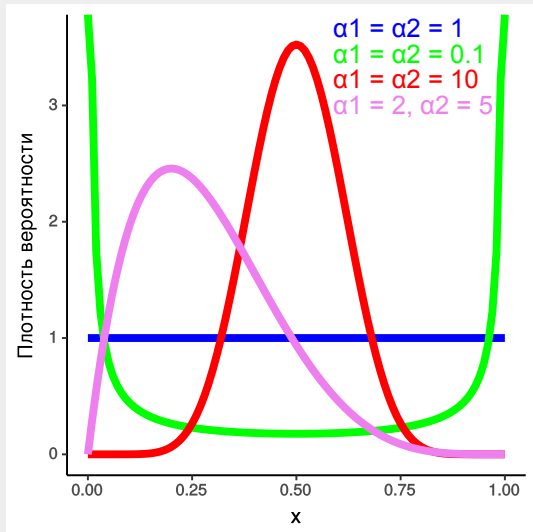


Гамма-распределение

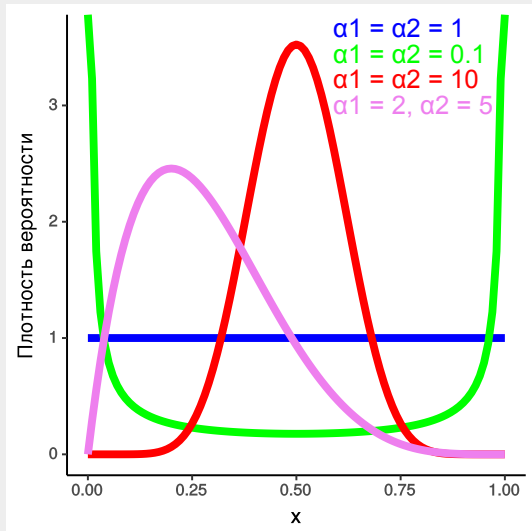


Сумма экспоненциальных распределений.
Длина ветвей, скорость эволюции, распределение
между позициями *etc.*

Бета-распределение и распределение Дирихле



Бета-распределение и распределение Дирихле



Любые доли (например, частота замен)

Программы для филогенетического анализа с помощью BI

Список программ



1. MrBayes



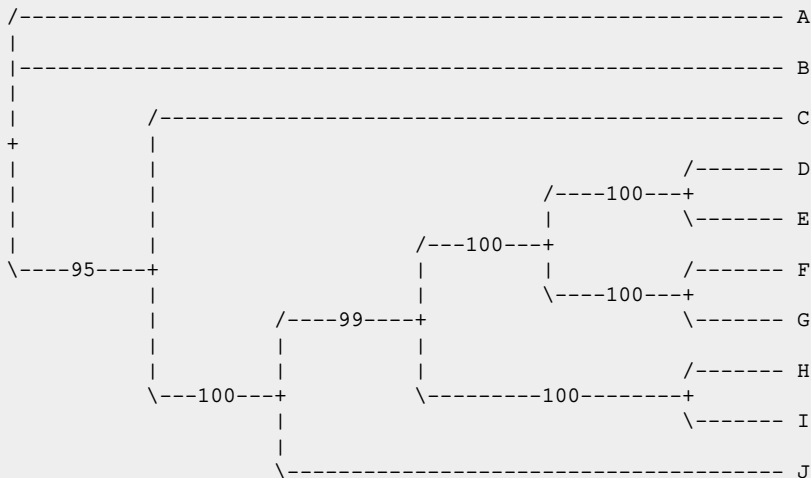
2. Beast

Пример вывода MrBayes

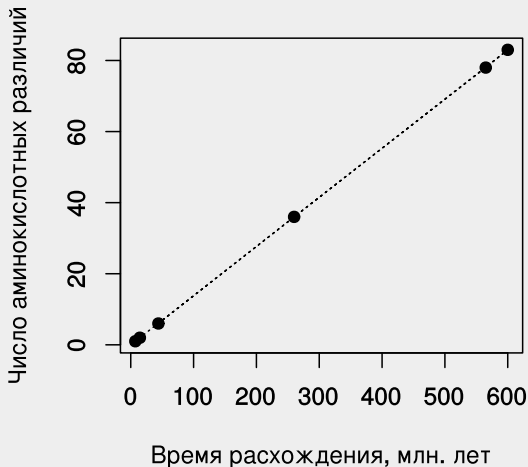
Суммирует деревья.

Около узлов — апостериорная вероятность.

Clade credibility values:



Гипотеза молекулярных часов (molecular clock)



По: Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic heterogeneity. In: Kasha M, Pullman B (eds). Horizons in biochemistry. Academic Press, New York, pp 189–225.

Модели распределения скорости накопления замен

- Constant-rates (global clock) \leq strict clock
(«строгие» часы)

Модели распределения скорости накопления замен

- ▶ Constant-rates (global clock) \leq strict clock («строгие» часы)
- ▶ Variable-rates \leq relaxed clock («ослабленные») + random (случайные)
 - ▶ Autocorrelated
 - ▶ Continuous: ACLN, ACG, AOUP, ACIP
 - ▶ Episodic: ACE, ACPP
 - ▶ Independent: **UCLN**, UCG, IGR, UCE
 - ▶ Local Clock: AHLC, RLC
 - ▶ Mixture: FMM, DPP

По Chen, M. H., Kuo, L., & Lewis, P. O. (Eds.). (2014). Bayesian Phylogenetics: methods, algorithms, and applications. Chapman and Hall/CRC.

Модели распределения возраста узлов

- ▶ Tree-Wide

- ▶ Generic: uniform, Dirichlet
- ▶ Biological
 - ▶ Population-Level: constant, exponential, logarithmic, skyline
 - ▶ Species-Level: Yule, birth-death, general birth-death

- ▶ Node-Specific

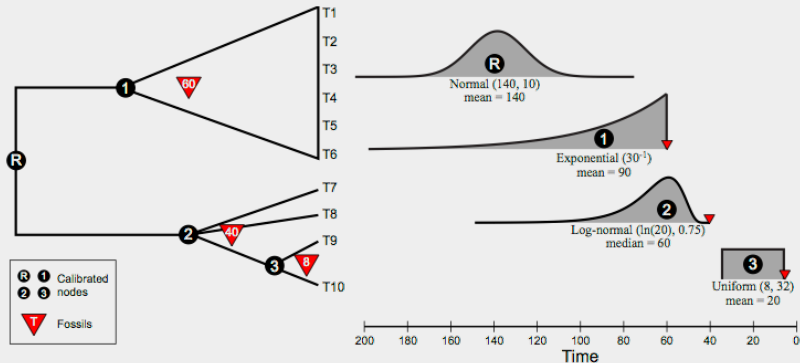
Chen, M. H., Kuo, L., & Lewis, P. O. (Eds.). (2014). Bayesian Phylogenetics: methods, algorithms, and applications. Chapman and Hall/CRC.

Данные для абсолютной калибровки:

1. биогеографические события;
2. серии образцов;
3. ископаемые останки.

Данные для абсолютной калибровки:

1. биогеографические события;
2. серии образцов;
3. ископаемые останки.



<https://sites.google.com/site/eeob563/computer-labs/lab-7>