Übungen Algorithmen der Bioinformatik II
Wintersemester 2005/06

Prof. Dr. Stefan Posch, Dr. Birgit Möller

Institut für Informatik
Universität Halle

Blatt 7

Aufgabe 7.1

a) Beweisen Sie den folgenden Satz und formulieren Sie ihn verbal. Sei

$$P(\vec{K}=\vec{k}\,|\,\vec{\lambda}) := \prod_{m=1}^{M} \frac{\lambda_m^{k_m}}{k_m!} e^{-\lambda_m} \ , \ K := K_1 + K_2, \text{ und } \lambda := \lambda_1 + \lambda_2.$$

Dann gilt

$$P(k, k_3, ..., k_M\,|\,\vec{\lambda}) := \sum_{k_1=0}^{k} P(k_1, k-k_1, k_3, ..., k_M\,|\,\vec{\lambda}) = \frac{\lambda^k}{k!} e^{-\lambda} \prod_{m=3}^{M} \frac{\lambda_m^{k_m}}{k_m!} e^{-\lambda_m}.$$

(Hinweis: Binomische Formel $(a+b)^n = \sum_{i=0}^{n} \frac{n!}{i!(n-i!)} a^i b^{n-i}$)

b) Definieren Sie $K = \sum_{i=1}^{M} K_i$ und verallgemeinern Sie den Beweis aus a).

c) Ein poissonscher DNA Generator feuere die Codons $1, 2, ..., 64$ mit den Raten $\lambda_1, \lambda_2, ..., \lambda_{64}$ statistisch unabhängig voneinander ab. Sei k_m die Anzahl der im Zeitintervall τ abgefeuerten Codons $m = 1, 2, ..., 64$. Wie lautet die Verbundwahrscheinlichkeitsverteilung $P(k_1, k_2, ..., k_{64}\,|\,\lambda_1, \lambda_2, ..., \lambda_{64})$?

O.B.d.A. seien $m = 1, 2, 3$ die Stopcodons, und $m = 4, 5, ..., 64$ die 61 aminosäurecodierenden Codons. Sei $k = k_1+k_2+k_3$ die Anzahl der Stopcodons, und $\ell = k_4+k_5+...+k_{64}$ die Anzahl der aminosäurecodierenden Codons. Wie lautet die Wahrscheinlichkeitsverteilung $P(k, \ell)$?

Aufgabe 7.2

a) Gegeben sei ein Poissonprozess, der die Zufallsvariablen k_{ij} (mit $i = 1, 2, ..., I$ und $j = 1, 2, ..., J$) mit der $I \times J$ dimensionalen Poissonverteilung

$$P(\mathbf{k}\,|\,\lambda) = \prod_{i=1}^{I} \prod_{j=1}^{J} \frac{\lambda_{ij}^{k_{ij}}}{k_{ij}!} e^{-\lambda_{ij}}$$

generiert, wobei $\mathbf{k}$ die Matrix der Zufallsvariablen $\{k_{ij}\}$ und λ die Matrix der Parameter $\{\lambda_{ij}\}$ des Poissonprozesses bezeichnet. Dieser Poissonprozess generiere L Datenpunkte (Matrizen) $\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_L$. Berechnen Sie den Maximum-Likelihood-Schätzer $\hat{\lambda}$ des Parameters λ.

b) Sei nun $\lambda_{ij} = \alpha_i \beta_i$ mit $\vec{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_I)$ und $\vec{\beta} = (\beta_1, \beta_2, ..., \beta_J)$, also

$$P(\mathbf{k}\,|\,\vec{\alpha}, \vec{\beta}) = \prod_{i=1}^{I} \prod_{j=1}^{J} \frac{(\alpha_i \beta_j)^{k_{ij}}}{k_{ij}!} e^{-\alpha_i \beta_j}$$

Berechnen Sie die Maximum-Likelihood-Schätzer $\hat{\vec{\alpha}}$ und $\hat{\vec{\beta}}$.

Aufgabe 7.3

Beweisen Sie, dass die zur Exponentialverteilung konjugierte Verteilung eine Gammaverteilung ist. Wie transformieren sich die Parameter α und λ?

Abgabe: 8.12.05