Algorytmika
Ćwiczenia 6

Adrian Herda
Politechnika Wrocławska

2025/05/28

Zadanie 27

Niech X będzie licznikiem morrisa. Licznik rozpoczyna z wartością 0 i dla każdej operacji inkremetowany jest z prawdopodobieństwem 2^{-X}. Estymator morrisa jest zdefiniowany jako:

$$\hat{n} = 2^{X_n} - 1$$
$$\hat{n} + 1 = 2^{X_n}$$

gdzie X_n jest licznikiem morrisa po n operacjach. Wartość oczekiwana tego estymatora jest równa:

$$E[\hat{n}] = n$$

Dowód przez indukcję:
1. Dla $n = 0$ mamy $E[\hat{n}] = 2^{X_0} - 1 = 2^0 - 1 = 0$.
2. Załóżmy, że dla dowolnego n mamy $E[\hat{n}] = n$.
3. Dla $n + 1$ mamy:

$$\begin{aligned}
E\left[\widehat{n+1}\right] &= E[2^{X_{n+1}} - 1] \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot E[2^{X_{n+1}} \mid X_n = i] - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot \left(\underbrace{\frac{1}{2^i} \cdot 2^{i+1}}_{\text{increment}} + \underbrace{\left(1 - \frac{1}{2^i}\right) \cdot 2^i}_{\text{no increment}} \right) - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i) \cdot (2^i + 1) - 1 \\
&= \sum_{i=1}^{\infty} P(X_n = i) 2^i + \sum_{i=1}^{\infty} P(X_n = i) - 1 \\
&= E[2^{X_n}] + 1 - 1 \\
&= E[\hat{n} + 1] = n + 1
\end{aligned}$$

∎

Wariancja

$$\begin{aligned}
E[2^{2X_n}] &= \sum_{i=1}^{\infty} P(X_n = i) \cdot 2^{2i} \\
&= \sum_{i=1}^{\infty} \left(P(X_{n-1} = i-1) \cdot \frac{1}{2^{i-1}} + P(X_{n-1} = i) \cdot \left(1 - \frac{1}{2^{i-1}}\right)\right) \cdot 2^{2i} \\
&= \sum_{i=1}^{\infty} 2^{i+1} P(X_{n-1} = i-1) + \sum_{i=1}^{\infty} 2^{2i} P(X_{n-1} = i) - \sum_{i=1}^{\infty} 2^{i} P(X_{n-1} = i) \\
&= 4 \sum_{i=1}^{\infty} 2^{i-1} P(X_{n-1} = i-1) + E[2^{2X_{n-1}}] + E[2^{X_{n-1}}] \\
&= 4E[2^{X_{n-1}}] + E[2^{2X_{n-1}}] + E[2^{X_{n-1}}] \\
&= 3E[2^{X_{n-1}}] + E[2^{2X_{n-1}}] \\
&= 3n + E[2^{2X_{n-1}}]
\end{aligned}$$

Ta formuła jest rekurencyjna, więc możemy j rozwin i otrzymamy ciąg arytmetyczny:

$$E[2^{2X_n}] = 1 + \sum_{i=1}^{n} 3i = 1 + \frac{3}{2} n(n+1)$$

A więc:

$$\begin{aligned}
\mathrm{Var}(\hat{n}) &= E\left[\left(2^{X_n} - 1\right)^2\right] - E[2^{X_n} - 1]^2 \\
&= E[2^{2X_n} - 2 \cdot 2^{X_n} + 1] - n^2 \\
&= E[2^{2X_n}] + 1 - 2E[2^{X_n}] - n^2 \\
&= 1 + \frac{3}{2} n(n+1) + 1 - (2E[2^{X_n}] - 2) - n^2 - 2 \\
&= \frac{3}{2} n^2 + \frac{3}{2} n - (2E[2^{X_n} - 1]) - n^2 \\
&= \frac{3}{2} n^2 + \frac{3}{2} n - 2n - n^2 \\
&= \frac{1}{2} n^2 - \frac{1}{2} n \\
&= \frac{1}{2} n(n-1)
\end{aligned}$$

Błąd standardowy

Błąd standardowy estymatora morrisa jest równy pierwiastkowi z wariancji podzielonej przez liczbę operacji:

$$\mathrm{SE}[\hat{n}] = \sqrt{\mathrm{Var}\left(\frac{\hat{n}}{n}\right)} = \sqrt{\frac{1}{2} \frac{n(n-1)}{n^2}} = \sqrt{\frac{n-1}{2n}}$$

Zadanie 28

Niech $h(x)$ będzie funkcją hashującą, która dla każdego x zwrca wartość z tablicy haszującej o długości m.

$$P(\text{kolizja}) = P\left(\exists_{i \neq j}\left(h(x_i) = h\left(x_j\right)\right)\right)$$

Dowód wzoru

Załóżmy że wstawione zostało już n elementów. To znaczy że pozstało $m - n$ miejsc w tablicy haszującej. Prawdopodobieństwo kolizji przy wstawianiu następnego elementu $x_{\{n+1\}}$ jest równe:

$$P(\text{kolizja}) = \frac{n}{m}$$

Przy wstawainiu kolejnch n elementów prawdopodobieństwo kolizji liczone jesst w następujący sposób:

$$\left.\begin{array}{l} P(\text{brak kolizji}) = \prod_{i=0}^{n-1} \frac{m-i}{m} \approx \exp\left(-\frac{n(n-1)}{2m}\right) \\ \frac{1}{2} \leq P(\text{kolizja}) \Longrightarrow \frac{1}{2} \geq P(\text{brak kolizji}) \end{array}\right\} \Longrightarrow \frac{1}{2} \approx \exp\left(-\frac{n(n-1)}{2m}\right)$$

$$\ln(2) \approx \frac{n(n-1)}{2m}$$
$$n(n-1) \approx 2\ln(2)m$$
$$n^2 - n \approx 2\ln(2)m$$

Dla dostatecznie dużych n możemy przyjąć, że

$$n^2 - n \approx n^2$$

więc:

$$n^2 \approx 2\ln(2)m$$
$$n \approx \sqrt{2\ln(2)m}$$

■

Wartość n dla $m = 2^{16}$

$$\begin{aligned} n &\approx \sqrt{2\ln(2)2^{16}} \\ &\approx \sqrt{2 * 0.693147 * 65536} \\ &\approx \sqrt{90852.18725} \\ &\approx 301.417 \end{aligned}$$

Zadanie 29