

Algorytmy probabilistyczne

Lista zadań nr 7

1. Do n urn losowo wrzucamy n kul (jednostajnie i niezależnie). Udowodnić, że z prawdopodobieństwem co najmniej $1 - 1/n$ zachodzi zdarzenie, że kul w każdej urnie jest co najwyżej $3 \ln n / \ln \ln n$, dla dostatecznie dużego n .
2. Udowodnić, że jeśli zmienna losowa Z ma rozkład Poissona z całkowitym parametrem μ spełniającym warunek $\mu \geq 1$, to $Pr(Z \geq \mu) \geq \frac{1}{2}$. W tym celu wykazać, że $Pr(Z = \mu + h) \geq Pr(Z = \mu - h - 1)$ dla $0 \leq h \leq \mu - 1$.
3. Niech $X_i^{(m)}$, $1 \leq i \leq n$, oznacza liczbę kul w i -tej urnie po losowym wrzuceniu m kul do n urn (jednostajnie i niezależnie). Niech $f: R^n \mapsto R$ będzie niemalejąca. Wykazać, że dla $k > m$ zachodzi: $E[f(X_1^{(m)}, \dots, X_n^{(m)})] \leq E[f(X_1^{(k)}, \dots, X_n^{(k)})]$.
4. Używając oznaczeń z poprzedniego zadania udowodnić, że jeśli A jest zdarzeniem zależnym od zmiennych $X_1^{(m)}, \dots, X_n^{(m)}$, którego prawdopodobieństwo jest niemalejącą funkcją m w modelu dokładnym, to $Pr[A]$ można oszacować od góry przez podwojone prawdopodobieństwo tego zdarzenia w modelu Poissona.
5. Wrzucamy losowo n kul do n urn. Niech X_i , $i = 1 \dots n$, będzie 0-1 zmienną losową oznaczającą, że i -ta urna pozostała pusta, a $X = \sum_{i=1}^n X_i$ określa liczbę pustych urn. Niech Y_i , $i = 1 \dots n$ będą niezależnymi 0-1 zmiennymi losowymi, takimi że $Pr(Y_i = 1) = (1 - 1/n)^n$ oraz niech $Y = \sum_{i=1}^n Y_i$. Wykazać, że: (1) $E[X_1 X_2 \dots X_k] \leq E[Y_1 Y_2 \dots Y_k]$ dla $k \geq 1$, oraz (2) $E[e^{tX}] \leq E[e^{tY}]$ dla każdego $t \geq 0$ i wyprowadzić z tego oszacowanie w stylu Chernoffa dla $Pr(X > (1 + \delta)E[X])$.
6. Załóżmy, że zbiory A i B zawierające po n elementów zostały wpisane do filtrów Blooma T_A i T_B rozmiaru m przy użyciu tych samych k funkcji haszujących. Filtr Blooma dla sumy zbiorów $A \cup B$ można wyznaczyć wykonując bitowy OR na elementach T_A i T_B . A co z różnicą zbiorów? Wyznaczyć oczekiwaną liczbę pozycji na których bity w T_A i T_B się różnią jako funkcję n , m , k i $|A \cap B|$. Czy można użyć tej wielkości do szacowania rozmiaru różnicy symetrycznej tych zbiorów?
7. Zamiana w filtrach Blooma pojedynczych bitów na b -bitowe liczniki pozwala zaimplementować także operację $Delete(x, S)$, gdy S jest zapisany w takim uogólnionym filtrze oraz $x \in S$. Podać szczegóły implementacji takiego uogólnienia. W praktyce często przyjmuje się $b = 4$. Oszacować prawdopodobieństwo wystąpienia błędu w operacji $Find(x, S)$ po wykonaniu t operacji wstawiania i usuwania na początkowo pustym S .