

1 Theoretical part

1.1 (due Jan 29, 2008) Babai (Finite Probability Spaces): 7, 11, 14, 16, 18, 21, 29.

1.2 (due Jan 29, 2008) CLRS: 5.4-1, 5.4-2, 5.4-3, 5.4-4, 5.4-5, 5.4-6, 5.4-7.

1.3 (due Jan 29, 2008) CLRS: C.3-1, C.3-2, C.3-4, C.3-7, C.3-9.

2 Applied part

1.4 (due Jan 29, 2008) Write a program which gets $\varepsilon, \delta, t \in (0, 1)$ on input and then it outputs N such that for every $p \in [t, 1]$ we have

$$P\left(\left|\frac{X_1 + \dots + X_N}{N} - p\right| \geq \varepsilon p\right) \leq \delta,$$

where X_1, \dots, X_N are independent, identically, distributed random variables with Bernoulli distribution with parameter p .

COMMENT 1: One can, of course, use Chernoff bound, however the goal here is to get a smaller value of N than what Chernoff bound would give.

COMMENT 2: The values of ε, δ, t are given as rational numbers. Make sure that your program always outputs a correct answer and is not a victim of imprecise floating point arithmetic. (You might find the GMP library useful <http://gmplib.org/> (you will definitely use it in later projects, so this is a good time to explore it)).

What values of N does your program give for:

- For $\varepsilon = 1/10, \delta = 1/10$ and $p = k/20$ for $k = 1, \dots, 19$.
- For $\varepsilon = 1/10, \delta = 1/1000$ and $p = k/20$ for $k = 1, \dots, 19$.
- For $\varepsilon = 1/1000, \delta = 1/10$ and $p = k/20$ for $k = 1, \dots, 19$.
- For $\varepsilon = 1/1000, \delta = 1/1000$ and $p = k/20$ for $k = 1, \dots, 19$.

Compare the values your program gives with what Chernoff bound suggests.

1.5 (due Jan 29, 2008) There is a value $p \in (0, 1)$ that you would like to estimate. You have access to a device which outputs independent random samples from a Bernoulli distribution with parameter p . Write a program which gets $\varepsilon, \delta \in (0, 1)$ on input, and then repeatedly asks for samples from the device and in the end it outputs Y such that

$$P(|Y - p| \geq \varepsilon p) \leq \delta.$$

Run your program with parameters $\varepsilon = 1/10, \delta = 1/10$ and for the following three settings of p : $p = 1/10, p = 1/100$, and $p = 1/1000$. How many samples did your algorithm use?