Probability Theory and Statistics
Lecture 17

Bence Csonka

Budapest University of Technology and Economics
csonkab@edu.bme.hu

November 17

Markov's Inequality

It may happen that we do not know the exact distribution of a random variable, but we can estimate its expectation, variance (e.g., based on sampling $\rightarrow$ see soon in the statistics part).

In the following we study inequalities that, based on expectation of a random variable, give an estimate for the probability that the random variable takes „extreme" (large/small) values.

Markov's Inequality

It may happen that we do not know the exact distribution of a random variable, but we can estimate its expectation, variance (e.g., based on sampling $\to$ see soon in the statistics part).

In the following we study inequalities that, based on expectation of a random variable, give an estimate for the probability that the random variable takes „extreme" (large/small) values.

Theorem

Let X be a nonnegative random variable. Then for every $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

(Proof)

Markov's Inequality

It may happen that we do not know the exact distribution of a random variable, but we can estimate its expectation, variance (e.g., based on sampling $\rightarrow$ see soon in the statistics part).

In the following we study inequalities that, based on expectation of a random variable, give an estimate for the probability that the random variable takes „extreme" (large/small) values.

Theorem

Let X be a nonnegative random variable. Then for every $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

(Proof)

True in the discrete, continuous (and general) case as well.

Chebyshev's Inequality

Theorem (Turbo Markov Inequality)

Let X be a random variable and let $g\colon \underbrace{\mathrm{Ran}(X)}_{\subseteq \mathbb{R}} \to [0,\infty)$ be continuous and strictly increasing. Then for every $a > \mathbb{E}(g(X))$,

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

Strict monotonicity is important—otherwise the first equality need not hold!

Chebyshev's Inequality

Theorem (Turbo Markov Inequality)

Let X be a random variable and let $g\colon \underbrace{\mathrm{Ran}(X)}_{\subseteq\mathbb{R}} \to [0,\infty)$ be continuous and *strictly increasing*. *Then for every $a > \mathbb{E}(g(X))$,*

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

Strict monotonicity is important—otherwise the first equality need not hold!

Theorem ((Chebyshev's inequality))

Let Y be a random variable. Then for every $a > 0$,

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) \leq \frac{\mathbb{D}^2(Y)}{a^2}.$$

Proof of Chebyshev's

Theorem ((Chebyshev's inequality))

Let Y be a random variable. Then for every $a > 0$,

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) \leq \frac{\mathbb{D}^2(Y)}{a^2}.$$

Proof: This is just the turbo Markov with the following choice:

- $X = |Y - \mathbb{E}(Y)|$,
- $g(x) = x^2$ (which is strictly increasing on $\mathrm{Ran}(X)$, i.e., on $[0, \infty)$).

By the turbo Markov and the definition of the variance,

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) = \mathbb{P}((Y - \mathbb{E}(Y))^2 \geq a^2) \leq \frac{\mathbb{E}((Y - \mathbb{E}(Y))^2)}{a^2} = \frac{\mathbb{D}^2(Y)}{a^2}.$$

Proof of Chebyshev's

Theorem ((Chebyshev's inequality))

Let Y be a random variable. Then for every $a > 0$,

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) \leq \frac{\mathbb{D}^2(Y)}{a^2}.$$

Proof: This is just the turbo Markov with the following choice:

- $X = |Y - \mathbb{E}(Y)|$,
- $g(x) = x^2$ (which is strictly increasing on $\mathrm{Ran}(X)$, i.e., on $[0, \infty)$).

Chebyshev's inequality works for any random variable (although if $\mathbb{E}(Y^2) = \infty$, it yields a trivial bound).

Chebyshev's Inequality: Examples

A given database server handles on average 50 requests per unit time. From experience, the standard deviation of the number of requests is 5. Give a lower bound on the probability that the number of requests in a unit time is more than 40 but less than 60.

This problem is a typical example of when we do not know the distribution of a random variable, only something about its moments (here, its standard deviation).

Chebyshev's inequality does not always give a very good bound; indeed, even when $\mathbb{E}(Y^2) < \infty$ it may give a trivial bound.

Weak Law of Large Numbers

We will see an application of Chebyshev's inequality. Let $\overline{X}_n := \frac{X_1 + \ldots + X_n}{n}$, where X_is are indepdendent identacilly distributed with $\mathbb{E}(X_i) = \mu$, $\mathbb{D}(X_i) = \sigma$.

Weak Law of Large Numbers

We will see an application of Chebyshev's inequality. Let $\overline{X}_n := \frac{X_1+\ldots+X_n}{n}$, where X_is are indepdendent identacilly distributed with $\mathbb{E}(X_i) = \mu$, $\mathbb{D}(X_i) = \sigma$.

Theorem (Weak Law of Large Numbers)

$$\lim_{n\to\infty} \left(|\overline{X}_n - \mu| \geq \varepsilon\right) = 0.$$

(Proof)

Weak Law of Large Numbers

We will see an application of Chebyshev's inequality. Let $\overline{X}_n := \frac{X_1 + \ldots + X_n}{n}$, where X_is are indepdendent identacilly distributed with $\mathbb{E}(X_i) = \mu$, $\mathbb{D}(X_i) = \sigma$.

Theorem (Weak Law of Large Numbers)

$$\lim_{n \to \infty} \left(|\overline{X}_n - \mu| \geq \varepsilon \right) = 0.$$

(Proof)

Theorem (Strong Law of Large Numbers)

$$\mathbb{P}(\lim_{n \to \infty} \overline{X}_n = \mu) = 1.$$

(We will not prove.)

Parameterized Chernoff Bound

Theorem

Let X be a random variable and let $g\colon \mathrm{Ran}(X) \to [0,\infty)$ be continuous and strictly increasing. Then for every $a > \mathbb{E}(g(X))$,

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

Compared to the choice $g(x) = x^2$, we can obtain a sharper bound by choosing a faster growing g:

Theorem (Parameterized Chernoff inequality)

Let X be a random variable. Then for every $a, t > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}\left(\mathrm{e}^{tX}\right)}{e^{ta}}.$$

Parameterized Chernoff Bound

Theorem

Let X be a random variable and let $g\colon \mathrm{Ran}(X) \to [0,\infty)$ be continuous and strictly increasing. Then for every $a > \mathbb{E}(g(X))$,

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

Theorem (Parameterized Chernoff inequality)

Let X be a random variable. Then for every $a, t > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(\mathrm{e}^{tX})}{e^{ta}}.$$

Proof: This is indeed the turbo Markov with $X = X$, $g(x) = \mathrm{e}^{tx}$.

Example for Parameterized Chernoff inequality

Let $X \sim \mathrm{Pois}(5)$. Give an upper bound for $\mathbb{P}(X \geq 10)$. $\rightarrow$ Using the parameterized Chernoff inequality and optimizing over t yields a much tighter upper bound than Chebyshev's inequality.)

Exercises

Exercise 1: Let X be a non-negative random variable such that $\mathbb{E}(X) = 5$. Use Markov's inequality to give an upper bound on $\mathbb{P}(X \geq 15)$.

Exercises

Exercise 1: Let X be a non-negative random variable such that $\mathbb{E}(X) = 5$. Use Markov's inequality to give an upper bound on $\mathbb{P}(X \geq 15)$.

Exercise 2: Let X be a random variable with mean $\mu = 100$ and variance $\sigma^2 = 25$. Use Chebyshev's inequality to find an upper bound on the probability that X differs from its mean by more than 10.

Exercises

Exercise 1: Let X be a non-negative random variable such that $\mathbb{E}(X) = 5$. Use Markov's inequality to give an upper bound on $\mathbb{P}(X \geq 15)$.

Exercise 2: Let X be a random variable with mean $\mu = 100$ and variance $\sigma^2 = 25$. Use Chebyshev's inequality to find an upper bound on the probability that X differs from its mean by more than 10.

Exercise 3: Let X be a random variable with $\mathbb{E}(X) = 50$ and $\mathrm{Var}(X) = 9$. Find an upper bound on $\mathbb{P}(|X - 50| \geq 6)$ using Chebyshev's inequality. Then compute the bound for $\mathbb{P}(|X - 50| \geq 3)$.