NMAI059 Probability and statistics 1
Class 10

Robert Šámal

Inequalities we know from the last time (and one more)

- Markov:
$$X \geq 0 \Rightarrow P(X \geq a\mathbb{E}(X)) \leq \frac{1}{a}$$
- Chebyshev
$$P(|X - \mathbb{E}(X)| \geq a\sigma_X) \leq \frac{1}{a^2}$$
- Chernoff ($\sigma_X = \sqrt{n}$)
$$X = \sum_{i=1}^{n} X_i, X_i = \pm 1 \Rightarrow P(|X - \mathbb{E}(X)| > a\sigma_X) \leq 2e^{-a^2/2}$$

Overview

Limit theorems – approximation

Statistics – an introduction

Strong law of large numbers

Theorem

Let $X_1, \ldots, X_n$ be i.i.d. with expectation μ and variance σ^2. Let $S_n = (X_1 + \cdots + X_n)/n$ be the sample mean. Then we have

$$\lim_{n\to\infty} S_n = \mu \quad \textit{almost surely (i.e. with probability 1).}$$

We say that sequence S_n converges to μ almost surely. and write $S_n \xrightarrow{a.s.} \mu$).

Monte Carlo integration

How to compute $\int_{x \in A} g(x)dx$?
In particular

$$g(x) = \begin{cases} 1 & \text{for } x \in S \\ 0 & \text{otherwise} \end{cases}$$

... area of a circle

Weak law of large numbers

Theorem

Let $X_1, \ldots, X_n$ *be i.i.d. with expectation* μ *and variance* σ^2*. Let* $S_n = (X_1 + \cdots + X_n)/n$ *be the sample mean. Then for every* $\varepsilon > 0$ *we have*

$$\lim_{n \to \infty} P(|S_n - \mu| > \varepsilon) = 0.$$

We say that sequence S_n *converges to* μ *in probability and write* $S_n \xrightarrow{P} \mu$*).*

Law of Large numbers $\rightarrow$ Central Limit Theorem

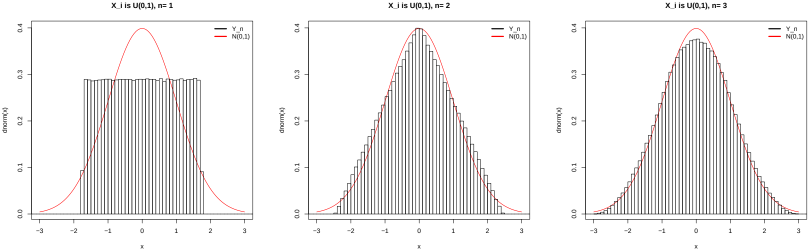
Central Limit Theorem

Theorem

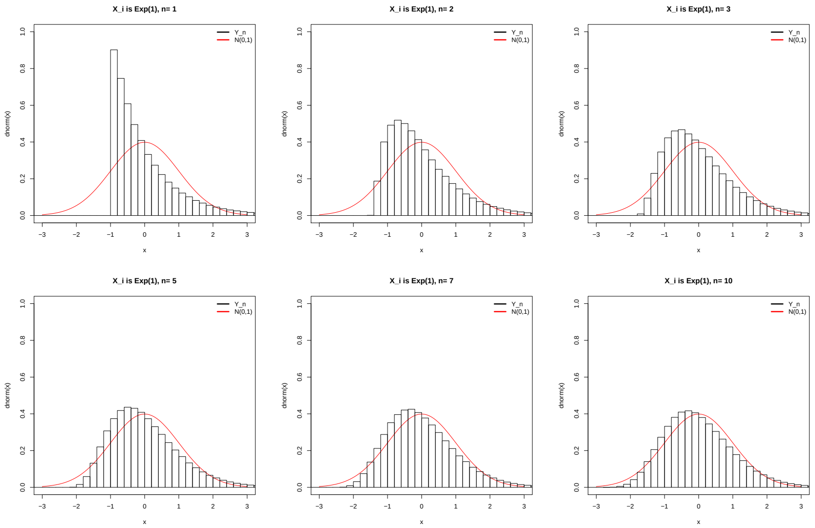
Let $X_1, \dots, X_n$ be i.i.d. with expectation μ and variance σ^2. Put $Y_n := ((X_1 + \cdots + X_n) - n\mu)/(\sqrt{n} \cdot \sigma)$.
Then $Y_n \xrightarrow{d} N(0,1)$. This means, that if F_n is the cdf of Y_n, then

$$\lim_{n \to \infty} F_n(x) = \Phi(x) \quad \textit{for every } x \in \mathbb{R}.$$

We say that the sequence Y_n converges to $N(0,1)$ in distribution.



CLT another illustration



Bonus: Moment generating function

Definition

For a random variable X we let

$$M_X(t) = \mathbb{E}\big(e^{tX}\big).$$

Function $M_X(t)$ is called the moment generating function.

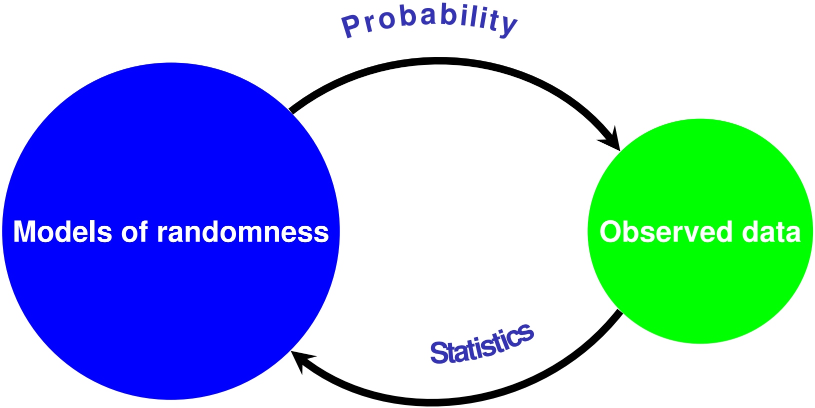
- $M_X(t) = \sum_{n=0}^{\infty} \mathbb{E}\big(X^n\big)\frac{t^n}{n!}$.
- $M_{Bern(p)}(t) = p \cdot e^t + (1-p)$.
- $M_{X+Y}(t) = M_X(t)M_Y(t)$, jsou-li X, Y n.n.v.
- $M_{Bin(n,p)} = (pe^t + 1 - p)^n$
- $M_{N(0,1)} = e^{t^2/2}$
- $M_{Exp(\lambda)} = \frac{1}{1-t/\lambda}$
- If $M_X(t) = M_Y(t)$ on $(-a, a)$ for some $a > 0$, then $X = Y$ a.s.

Overview

Limit theorems – approximation

Statistics – an introduction

Lecture overview



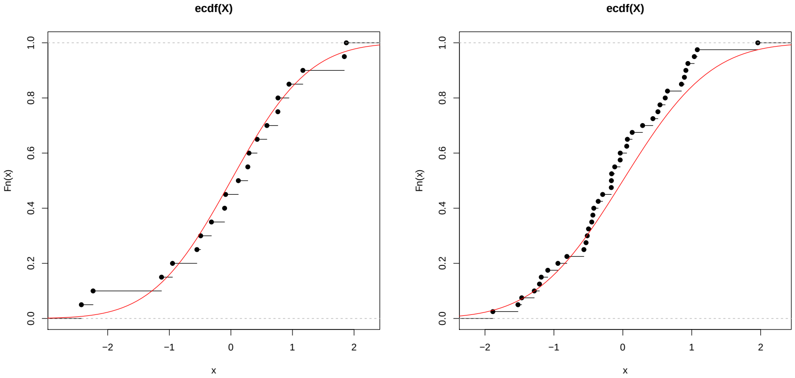
1st illustration – number of left-handed people

2nd illustration – running time of a program

- $X_1, \ldots, X_n \sim F$ i.i.d., F is their CDF
- **Definition:** *Empirical CDF* is defined by

$$\widehat{F}_n(x) = \frac{\sum_{i=1}^{n} I(X_i \leq x)}{n},$$

where $I(X_i \leq x) = 1$ if $X_i \leq x$ and 0 otherwise.



Empirical CDF – properties

Theorem

For a fixed x

- $\mathbb{E}(\widehat{F}_n(x)) = F(x)$
- $var(\widehat{F}_n(x)) = \frac{F(x)(1-F(x))}{n}$
- $\widehat{F}_n(x)$ *converges to* $F(x)$ *in probability,* $\widehat{F}_n(x) \xrightarrow{P} F(x)$.

Důkaz.

Weak law of large numbers.

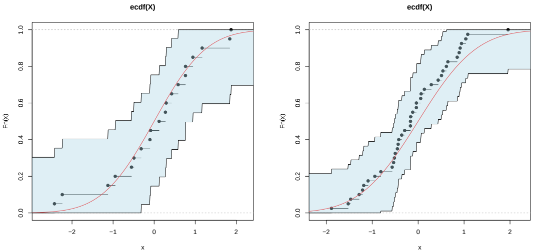
Note that $n\widehat{F}(x) \sim Bin(n, F(x))$ □

Empirical CDF – Dvoretzky-Kiefer-Wolfowitz (DKW)

Theorem

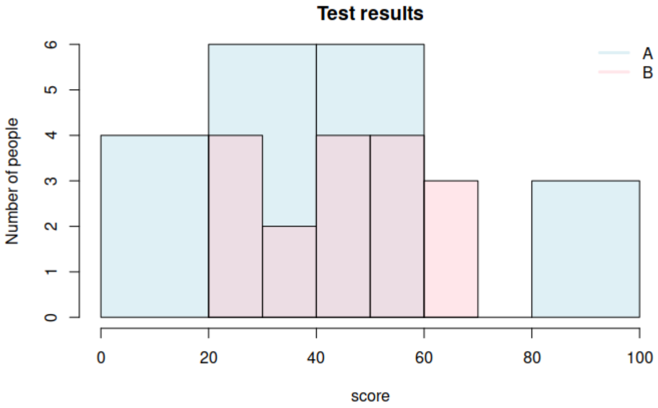
Let $X_1, \ldots, X_n \sim F$ *be i.i.d., let* $\widehat{F}_n$ *be their empirical CDF. Let* $\mathbb{E}(X_i)$ *be finite. Choose* $\alpha \in (0,1)$ *and let* $\varepsilon = \sqrt{\frac{1}{2n} \log \frac{2}{\alpha}}$*. The we have*

$$P(\widehat{F}_n(x) - \varepsilon \leq F(x) \leq \widehat{F}_n(x) + \varepsilon) \geq 1 - \alpha.$$



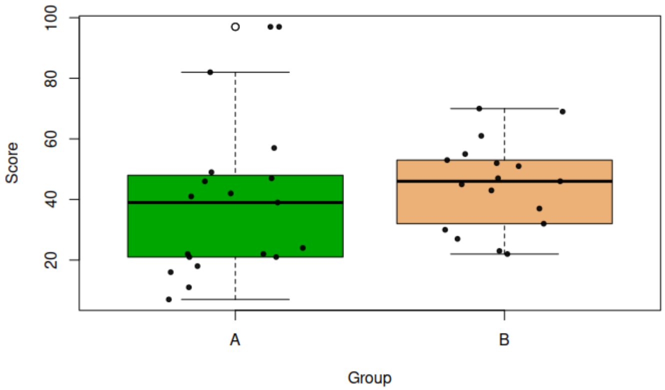
Intro – exploratory data analysis

- we collect data (and pay attention to systematic errors – independence, bias, . . .)
- we make various tables
- any appropriate charts: histogram, boxplot, etc.



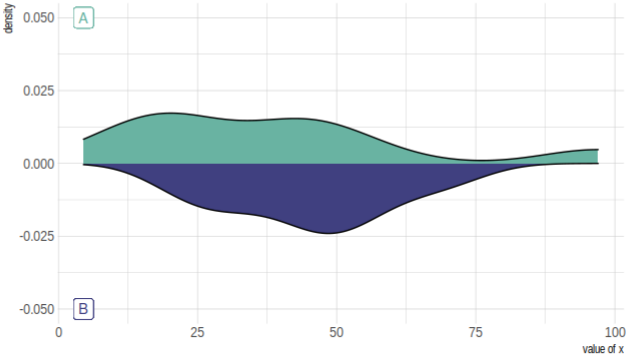
Intro – exploratory data analysis

- we collect data (and pay attention to systematic errors – independence, bias, . . .)
- we make various tables
- any appropriate charts: histogram, boxplot, etc.



Intro – exploratory data analysis

- we collect data (and pay attention to systematic errors – independence, bias, . . .)
- we make various tables
- any appropriate charts: histogram, boxplot, etc.



Goals of confirmatory data analysis

- point estimates
- interval estimates
- hypothesis testing
- (linear) regression

Examples:

- We assume human height follows $N(\mu, \sigma^2)$. What are μ and σ?
- Is our coin/dice fair?
- Is a medical treatment beneficial?
- Is new version of a program faster?
- How does running time depend on size of input?