Measurement Methods, Modelling and Processing of Experimental Data - NEVF503

Z. Němeček: November 20, December 4, 2023 13:10-16:10 CEST

1. Principles of modern measuring methods

2. Signal processing

O. Santolík: October 9, October 23, November 6, 2023 13:10-16:10 CEST

ondrej.santolik@mff.cuni.cz http://os.matfyz.cz/EVF503

3. Statistical methods of data processing

Probability, random vectors, pseudorandom sequences, bias and variance of estimates, correlation coefficients, parametric methods.

4. Data modelling

Interpolation, maximum likelihood methods, general linear least squares method, singular values decomposition (SVD): theory and examples, non-linear least squares methods, confidence intervals, Golay-Savitzky filters, splines.

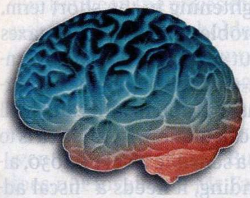
5. Random processes

Mean value, correlation, stationarity ergodicity, convolution, power spectra, multidimensional spectral analysis, wavelet analysis.

Motivation:

How science goes wrong

Scientific research has changed the world. Now it needs to change itself



A SIMPLE idea underpins science: “trust, but verify”. Results should always be subject to challenge from experiment. That simple but powerful idea has generated a vast body of knowledge. Since its birth in the 17th century, modern science has changed the world beyond recognition, and overwhelmingly for the better.

But success can breed complacency. Modern scientists are doing too much trusting and not enough verifying—to the detriment of the whole of science, and of humanity.

Too many of the findings that fill the academic ether are the result of shoddy experiments or poor analysis (see pages 21-24). A rule of thumb among biotechnology venture-capitalists is that half of published research cannot be replicated. Even that may be optimistic. Last year researchers at one biotech firm, Amgen, found they could reproduce just six of 53 “landmark” studies in cancer research. Earlier, a group at Bayer, a drug company, managed to repeat just a quarter of 67 similarly important papers. A leading computer scientist frets that three-quarters of papers in his subfield are bunk. In 2000-10 roughly 80,000 patients took part in clinical trials based on research that was later retracted because of mistakes or improprieties.

What a load of rubbish

Even when flawed research does not put people's lives at risk—and much of it is too far from the market to do so—it squanders money and the efforts of some of the world's best minds. The opportunity costs of stymied progress are hard to quantify, but they are likely to be vast. And they could be rising.

One reason is the competitiveness of science. In the 1950s, when modern academic research took shape after its successes in the second world war, it was still a rarefied pastime. The entire club of scientists numbered a few hundred thousand. As their ranks have swelled, to 6m-7m active researchers on the latest reckoning, scientists have lost their taste for self-policing and quality control. The obligation to “publish or perish” has come to rule over academic life. Competition for jobs is cut-throat. Full professors in America earned on average $135,000 in 2012—more than judges did. Every year six freshly minted PHDs vie for every academic post. Nowadays verification (the replication of other people's results) does little to advance a researcher's career. And without verification, dubious findings live on to mislead.

Careerism also encourages exaggeration and the cherry-picking of results. In order to safeguard their exclusivity, the leading journals impose high rejection rates: in excess of 90% of submitted manuscripts. The most striking findings have the greatest chance of making it onto the page. Little wonder that one in three researchers knows of a colleague who has pepped up a paper by, say, excluding inconvenient data from results “based on a gut feeling”. And as more research teams around the world work on a problem, the odds shorten that at least one will fall prey to an honest confusion between the sweet signal of a genuine discovery and a freak of the statistical noise. Such spurious correlations are often recorded in journals eager for startling papers. If they touch on drinking wine, going senile or letting children play video games, they may well command the front pages of newspapers, too.

Conversely, failures to prove a hypothesis are rarely even offered for publication, let alone accepted. “Negative results” now account for only 14% of published papers, down from 30% in 1990. Yet knowing what is false is as important to science as knowing what is true. The failure to report failures means that researchers waste money and effort exploring blind alleys already investigated by other scientists.

The hallowed process of peer review is not all it is cracked up to be, either. When a prominent medical journal ran research past other experts in the field, it found that most of the reviewers failed to spot mistakes it had deliberately inserted into papers, even after being told they were being tested.

If it's broke, fix it

All this makes a shaky foundation for an enterprise dedicated to discovering the truth about the world. What might be done to shore it up? One priority should be for all disciplines to follow the example of those that have done most to tighten standards. A start would be getting to grips with statistics, especially in the growing number of fields that sift through untold oodles of data looking for patterns. Geneticists have done this and turned an early torrent of specious results from genome sequencing into a trickle of truly significant ones.

Ideally, research protocols should be registered in advance and monitored in virtual notebooks. This would curb the temptation to fiddle with the experiment's design midstream so as to make the results look more substantial than they are. (It is already meant to happen in clinical trials of drugs, but compliance is patchy.) Where possible, trial data also should be open for other researchers to inspect and test.

The most enlightened journals are already becoming less averse to humdrum papers. Some government funding agencies, including America's National Institutes of Health, which dish out $30 billion on research each year, are working out how best to encourage replication. And growing numbers of scientists, especially young ones, understand statistics. But these trends need to go much further. Journals should allocate space for “uninteresting” work, and grant-givers should set aside money to pay for it. Peer review should be tightened—or perhaps dispensed with altogether, in favour of post-publication evaluation in the form of appended comments. That system has worked well in recent years in physics and mathematics. Lastly, policymakers should ensure that institutions using public money also respect the rules.

Science still commands enormous—if sometimes bemused—respect. But its privileged status is founded on the capacity to be right most of the time and to correct its mistakes when it gets things wrong. And it is not as if the universe is short of genuine mysteries to keep generations of scientists hard at work. The false trails laid down by shoddy research are an unforgivable barrier to understanding. ■

1. Classical probability

of an event ***A*** is
the number of cases favorable to the event (number of elementary events composing ***A***),
divided by the number of all possible cases (number of elements of the sample space).

$$P(A) = \frac{N_A}{N}$$

where each elementary event
is a single element of the sample space (set of elementary events)
and corresponds to one of the mutually exclusive outcomes
with equal probabilities.

- **Luca Pacioli** (Franciscan, mathematician, friend of Leonardo da Vinci, father of double-entry system): **Summa de arithmetica, geometrica, proportioni et proportionalità (1494).**
  - Game of chance with two players *A* and *B* who have equal chances of winning each round.
  - The players contribute equally to a prize pot, and agree in advance that the first player to have won 6 rounds will collect the entire prize.
  - The game is interrupted by external circumstances before either player has achieved victory: *A* has won 5 rounds, *B* has won 3 rounds. How does one then divide the pot fairly? Pacioli says: *A* gets 5/8, *B* gets 3/8

Girolamo Cardano: Liber de Ludo Aleae (1564?). The 16th century treatment of the probability calculus. It could be also viewed as a gambling manual. Notion of justice, based on Aristotle's rule for a just act: “...there is one general rule, namely, that we should consider the whole circuit (note: Cardano’s term for ‘sample space’), and the number of those casts which represent in how many ways the favorable result can occur, and compare to that number to the remainder of the circuit, and according to that proportion should the mutual wagers be laid so that one may contend on equal terms.“

Discussion of Blaise Pascal and Pierre de Fermat (1654) **: The 17th century beginnings of modern probability theory**
Problem reposed by a French writer Antoine Gombaud (Chevalier de Méré):

- Game of chance with two players *A* and *B* who have equal chances of winning each round.
- The players contribute equally to a prize pot, and agree in advance that the first player to have won 6 rounds will collect the entire prize.
- The game is interrupted by external circumstances before either player has achieved victory: *A* has won 5 rounds, *B* has won 3 rounds. How does one then divide the pot fairly? Pacioli said: *A* gets 5/8, *B* gets 3/8

Pierre de Fermat 1601 - 1665

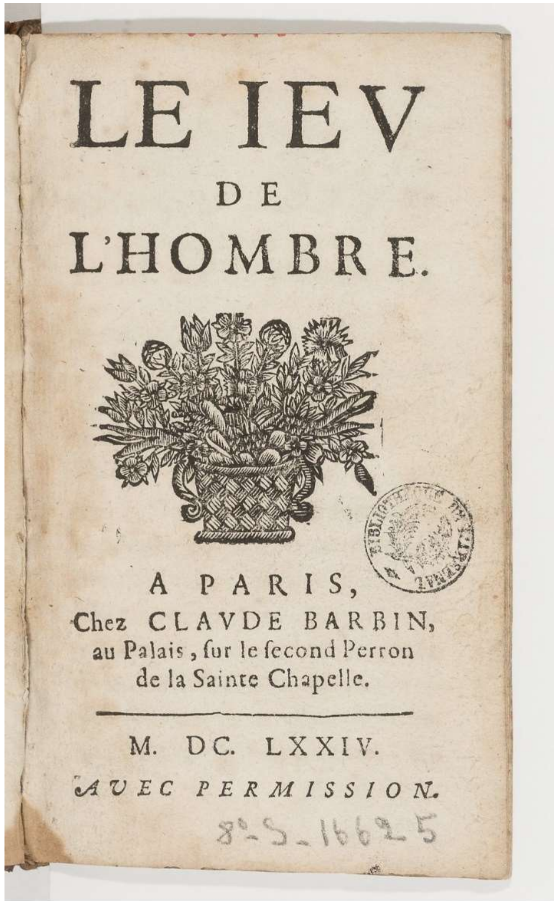
Fair solution:

- The game continues by 3 hypothetical rounds which have 2x2x2 = 8 possible outcomes.
- Only one of them is favorable to the event that *B* wins
- Therefore, fair distribution of funds from the pot is:

1/8 for *B*, 7/8 for *A*

Blaise Pascal 1623 - 1662

Another problem posed to Pascal by Chevalier de Méré:



- We need at least 4 rolls to obtain >50% probability of getting at least one "6" for a single 6-sided die:

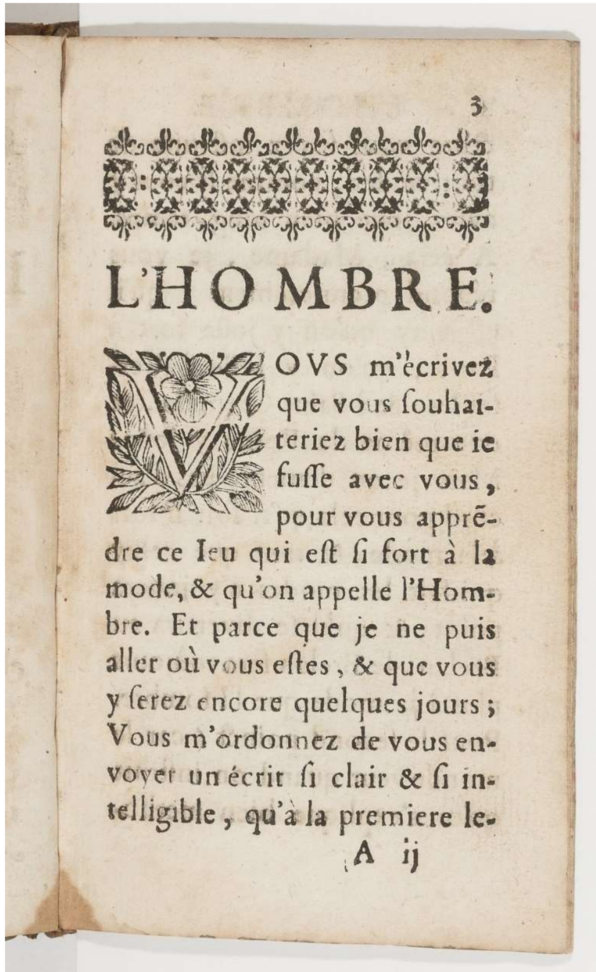
$$1-\left(\frac{5}{6}\right)^4 \approx 0.5177,$$

- How many rolls do we need to obtain >50% probability of getting at least one "6, 6" for two 6-sided dies:

  - intuition: 6 times more = 24
  - Chevalier de Méré guessed that intuition is misleading in this case. He posed the question to Pascal, who solved the problem and proved de Méré correct:

$$1-\left(\frac{35}{36}\right)^{24} \approx 0.4914.$$

$$1-\left(\frac{35}{36}\right)^{25} \approx 0.5055.$$



Disadvantage of classical probability: too special

2. A. N. Kolmogorov: Grundbegriffe der Wahrscheinlichkeitsrechnung (Springer, Berlin, 1933)

- $\Omega \ldots$ Sample space (set of all elementary events)
- **Event:** subset $A \subseteq \Omega$ (not all possible subsets are necessarily events)
- $\mathcal{B} \ldots$ Event space: set of all events, system of subsets of Ω.

Assumptions:

1. $\Omega \in \mathcal{B}$. *(the whole set of elementary events is an event)*
2. $A \in \mathcal{B} \Longrightarrow \Omega \backslash A$ ...complement $\in \mathcal{B}$. Then: $\emptyset \in \mathcal{B}$. *(empty set is an event)*
3. $A_1, A_2, \ldots A_n \in \mathcal{B}$ ... finite or at least countable system of events

$$\Longrightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{B}.$$

(union of events is an event)

Therefore (De Morgan's laws) :

$$\bigcap_{i=1}^{n} A_i \in \mathcal{B}.$$

(intersection of events is an event)

Andrey Nikolaevich Kolmogorov (1903-1983)

DEFINITION: $\mathcal{B}$ is sigma-algebra

Kolmogorov probability axioms (1933)

1. The probability of an event $A \in \mathcal{B}$ is a non-negative real number $P(A)$

2. The probability of the sample space (which is an event) is 1: $P(\Omega) = 1$
   In other words: the probability that at least one of the elementary events in the entire sample space will occur is 1.

3. If $A_i, i = 1 \ldots n$ is a countable system of disjoint sets (with empty intersections) representing mutually exclusive events - none of events can occur together with either of other events:

$$A_i \cap A_j = \emptyset, \forall i \neq j, i, j = 1 \ldots n$$

then

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

Some consequences of Kolmogorov probability axioms:

1. The probability of an event represented by empty set (never occurs) is $P(\emptyset) = 0$

2. If *A* is a subset of, or equal to *B*, then the probability of *A* is less than, or equal to the probability of *B*: $\text{if} \quad A \subseteq B \quad \text{then} \quad P(A) \leq P(B).$

   Proof: $P(A) + P(B \setminus A) = P(B)$

3. The probability of a complement to the sample space (the probability that the event will not happen: $P(\Omega \setminus A) = 1 - P(A)$

4. Sum rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B).$

The probability of a union of *A* and *B* (probability that an event *A* **or** event *B* will happen) is the sum of the probability of an event *A* and the probability of an event *B*, minus the probability of an intersection of *A* and *B* (probability that both *A* **and** *B* happen).

3. Conditional probability

Let $\{\Omega, \mathcal{B}, P\}$ be a measure space: Ω is a set, $\mathcal{B}$ is sigma-algebra on the set Ω, P is a measure on $(\Omega, \mathcal{B})$
probability space, when P is probability

Then, we can define the **conditional probability** as the probability of an event *A* occurring, given that another event *B* has already occurred (by assumption or by evidence):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

"the probability of *A* under the condition *B*"
or "the conditional probability of *A* given *B*"

Properties:

1. $0 \leq P(A|B) \leq 1.$
2. A_i ...system of disjoint events $A_i \cap A_j = \emptyset, \forall i \neq j$:

$$\Longrightarrow P(\bigcup_{i=1}^{n} A_i|B) = \sum_{i=1}^{n} P(A_i|B).$$

3. Occurrence of event *B* implies occurrence of event *A*:

$$B \subset A \Longrightarrow P(A|B) = 1.$$

4. Probability that both *A* **and** *B* happen (probability of intersection of *A* and *B*):

$$P(A \cap B) = P(A|B)\, P(B).$$

4. Independent events

- Definition: Events *A* and *B* are **independent** if $P(A \cap B) = P(A)\ P(B)$
- Consequence for the conditional probability of *A* given $P(A|B) = P(A)$.
- Different from **mutually exclusive** events for which the intersection, and $P(A|B)$ are $= 0$

5. Total probability

- $B_1, \ldots, B_n$ are disjoint events fulfilling $\Omega = \bigcup_{i=1}^{n} B_i$, it means that they form a **complete system** of events $(P(\bigcup B_i) = \sum P(B_i) = 1)$
- Then, for an event *A*: $$P(A) = \sum_{i=1}^{n} P(B_i)\ P(A|B_i).$$
- Proof: $$A = \bigcup_{i=1}^{n} (B_i \cap A)$$

6. Bayes' theorem (a.k.a. Bayes rule, Bayes' law)

Let $B_1, \ldots, B_n$ be a **complete system** of events (hypotheses)

Then, for an event *A* (outcome of an experiment):

$$P(B_k|A) = \frac{P(A|B_k)\ P(B_k)}{\sum_{i=1}^{n} P(A|B_i)\ P(B_i)}$$

$P(B_k|A)$ ... **posterior probabilities** of hypotheses $B_1, \ldots, B_n$

$P(B_k)$ ... **prior probabilities** of hypotheses $B_1, \ldots, B_n$

7. Random variable

Random variable ξ is a real function defined on the set of elementary events (sample set) mapping $\Omega \rightarrow \mathbb{R}$ such, that sets of all $\omega \in \Omega$ fulfilling $\xi(\omega) \leq a$ **are events** $\forall a \in \mathbb{R}$

$$\{\omega : \xi(\omega) \leq a\} \in \mathcal{B} \quad \forall a \in \mathbb{R},$$

and hence have probability.

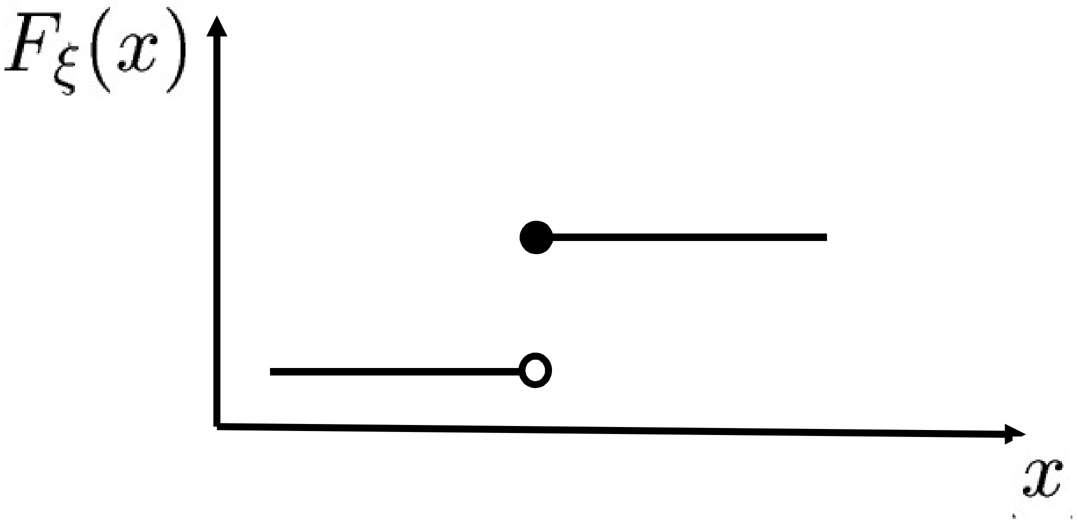
8. Distribution function of a random variable (a.k.a. cumulative distribution function, CDF)

Distribution function is a real function of a real variable, which returns **probability** that $\xi \leq x$

$$F_\xi(x) = P(\xi \leq x)$$

Properties of the distribution function

- non-decreasing
- $\lim_{x \to -\infty} F_\xi(x) = 0;$
- $\lim_{x \to \infty} F_\xi(x) = 1;$
- right-continuous
- countable number of discontinuities



for discrete random variables

Example: random variable $\xi =$ sum of points on two dice

- $x_i = 2, 3, \ldots, 6, 7, 8, \ldots, 12.$
- $P(\xi = x_i) = \frac{1}{36}, \frac{2}{36}, \ldots, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \ldots, \frac{1}{36}.$
- $F_\xi(x_i) = \frac{1}{36}, \frac{3}{36}, \ldots, \frac{15}{36}, \frac{21}{36}, \frac{26}{36}, \ldots, \frac{36}{36}.$

9. Probability density function (PDF), probability mass function (PMF)

The **probability density function** $f(u)$ is defined for a continuous random variable ξ with a cumulative distribution F_ξ :

$$F_\xi(x) = \int\limits_{-\infty}^{x} f(u)du,$$

for a discrete random variable ξ : **probability mass function** (a.k.a. discrete density function) a function giving the probability that a discrete random variable is exactly equal to some value

$$F_\xi(x) = \sum_{i, x_i \leq x} P(\xi = x_i)$$

10. Expected value, expectation, mathematical hope

Idea: 17th century, Pascal, de Fermat during the discussion on the division of stakes. **Christiaan Huygens**, 1657, *De ratiociniis in ludo aleæ*: „…If I expect a or b, and have an equal chance of gaining them, my Expectation is worth (a+b)/2.“

Pierre-Simon Laplace, 1814, *Théorie analytique des probabilités:* …this advantage in the theory of chance is the product of the sum hoped for by the probability of obtaining it… We will call this advantage *mathematical hope*.“

Discrete random variable:

$$\langle \xi \rangle = E(\xi) = \sum_i x_i P(\xi = x_i).$$

Continuous random variable:

$$\langle \xi \rangle = E(\xi) = \int_{-\infty}^{\infty} u\ f(u)\ du.$$

Pierre-Simon, marquis de Laplace
1749 – 1827

Christiaan Huygens
1629 –1695

Expected value of a real function of a random variable

Discr. r.v.: $E\Big(\phi(\xi)\Big) = \sum_i \phi(x_i)\ P(\xi = x_i).$ Cont. r.v.: $E\Big(\phi(\xi)\Big) = \int_{-\infty}^{\infty} \phi(u)\ f(u)\ du.$

Moments

*k*th moment: $\phi(\xi) = \xi^k,$ $m_k(\xi) = E(\xi^k) = \int_{-\infty}^{\infty} u^k\ f(u)\ du.$ Note: $m_1(\xi) = E(\xi)$

Central moments

*k*th central moment: $M_k(\xi) = E\Big([\xi - E(\xi)]^k\Big) = \int_{-\infty}^{\infty} [u - E(\xi)]^k\ f(u)\ du.$

11. Variance as a measure of dispersion of a random variable

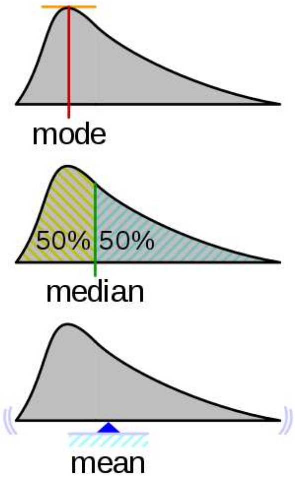
2nd central moment, square of the standard deviation σ

$$\mathrm{var}(\xi) = D(\xi) = \sigma^2(\xi) = M_2(\xi) = \int_{-\infty}^{\infty} [u - E(\xi)]^2\ f(u)\, du;$$

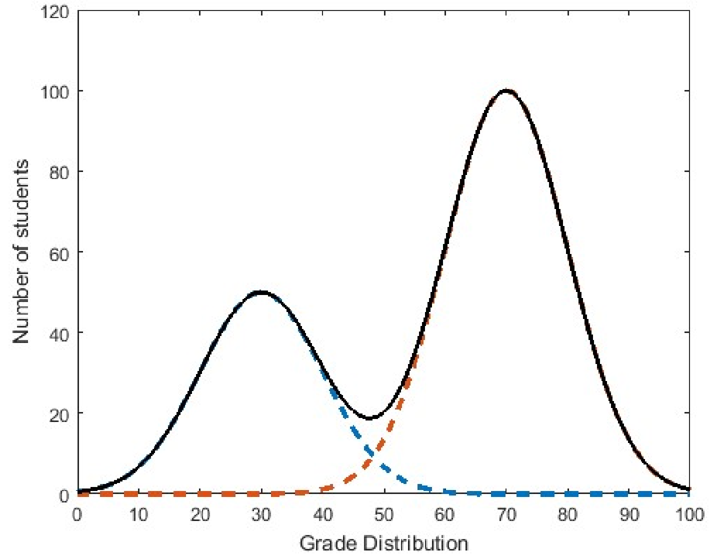
$$= \sum [x_i - E(\xi)]^2\, P(\xi = x_i)$$ for discrete random var.

12. Quantiles (percentiles, decils, quartiles)

x_p is a p-quantile of a random variable ξ : the cumulative distribution function $F(x_p) = p$



- $x_{0.5}$ ... **Median $= M$**
- $x_{0.25(0.75)}$ ... **lower (upper) quartile**
- $x_{\ell/10}$ ... ℓ-**th decile**
- $x_{\ell/100}$ ... ℓ-**th percentile**



13. Mode

value x_m for which the probability density function $f(x_m)$ has a local maximum
(more local maxima: bimodal or multimodal distribution)

Examples of discrete distributions

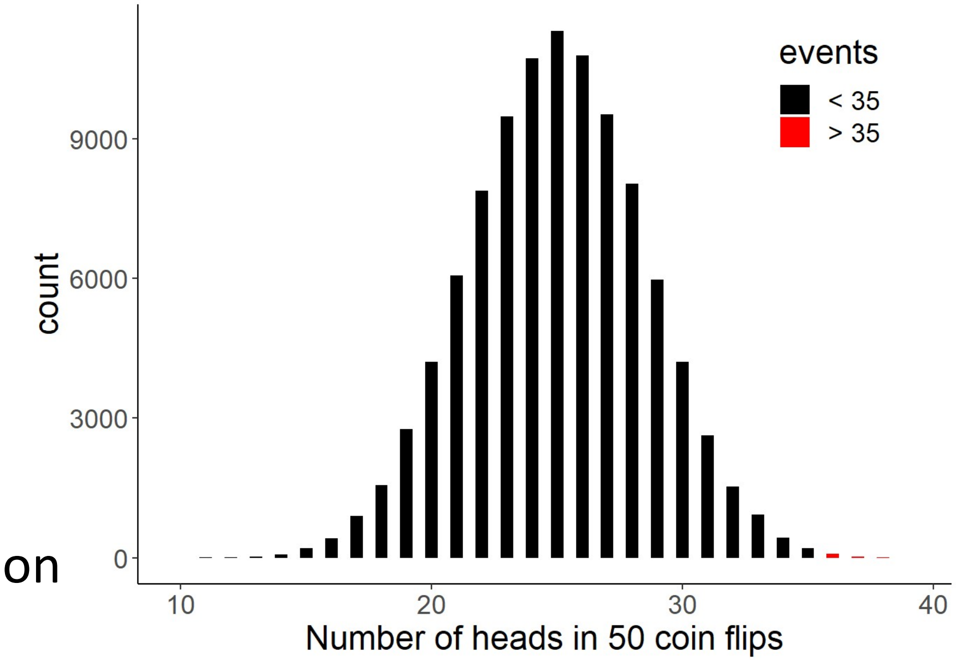
14. Binomial distribution

Discrete random variable (ξ): number of occurrences of an event in n independent experiments; occurrence probability of this event is equal to π in each particular experiment.

Probability mass function for $\xi = x, \quad x = 0 \ldots n$:

$$P(\xi = x) \quad = \quad \binom{n}{x} \pi^x \, (1-\pi)^{n-x}.$$

- Expected value : $E(\xi) = n\pi;$
- Variance: $\mathrm{var}(\xi) = n\pi(1-\pi).$



Binomial distribution converges to a normal distribution for large n

15. Poisson distribution

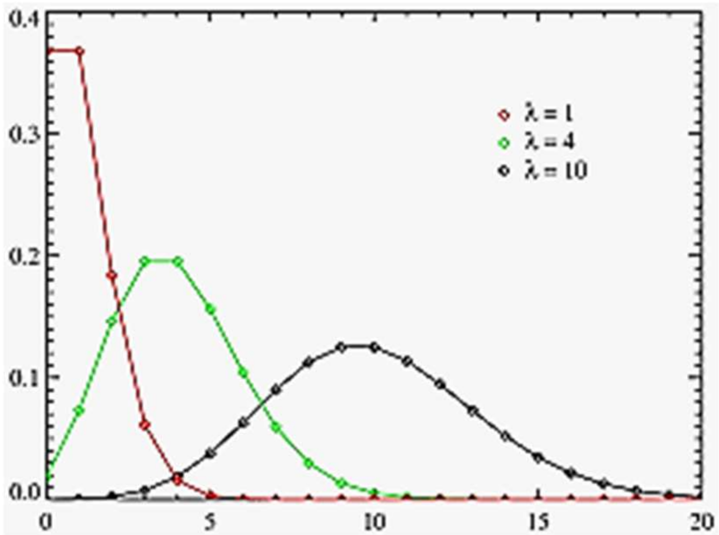
Limit of the binomial distribution for a large number of independent experiments when the probability of success in each particular experiment decreases as λ/n $(n \to \infty,\ n\pi \to \lambda)$.

Siméon Denis Poisson
1781-1840

Probability mass function:

$$P(\xi = x) \quad = \quad \exp(-\lambda)\frac{\lambda^x}{x!}$$

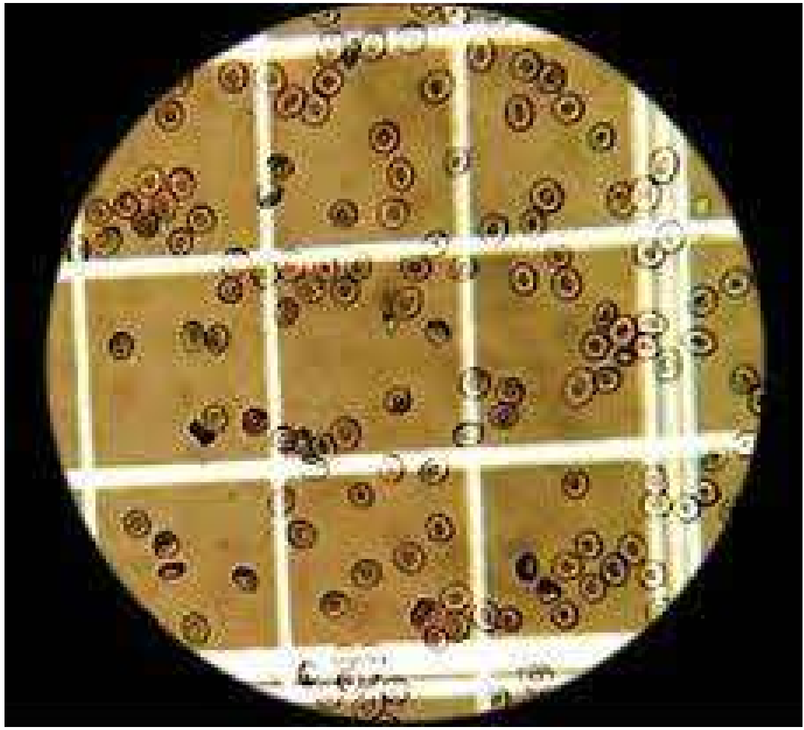
- **Expected value** : $E(\xi) = \lambda;$
- **Variance**: $\mathrm{var}(\xi) = \lambda \quad \longrightarrow \quad \sigma(\xi) = \sqrt{\lambda};$
- **Relative error** : $\sigma(\xi)/E(\xi) = 1/\sqrt{\lambda};$



Property of the **Poisson distribution**: if random variables $\xi_1, \xi_2 \ldots \xi_k$ have Poisson distributions with $\lambda_1, \lambda_2 \ldots \lambda_k$, then $\sum \xi_i$ has a Poisson distribution with $\sum \lambda_i$.

Examples:

- number of random values in a given interval,
- number of typos in a text,
- number of independent events occurring in a time interval τ: $\lambda = I\tau$
  - Radioactive decay counts
  - Large meteorites hitting the moon
- number of randomly placed points occurring in a predefined area (blood cells in a microscope)



Examples of continuous distributions

16. Uniform distribution

The probability density function is a constant on an interval $\langle a, b \rangle$:

$$f(u) = 0 \quad \text{pro} \quad (u < a) \vee (u > b); \qquad f(u) = \frac{1}{b-a} \quad \text{pro} \quad (u \geq a) \wedge (u \leq b).$$

Cumulative distribution function:

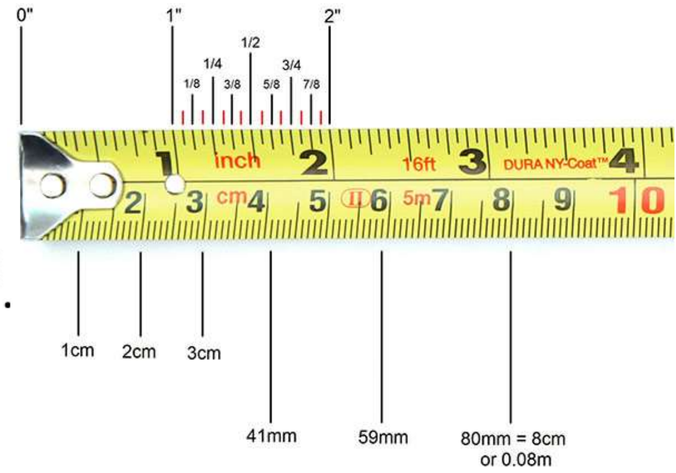
$F(x) = 0$ for $x \leq a$ $\qquad$ $F(x) = 1$ for $x \geq b;$ $\qquad$ $F(x) = \dfrac{x-a}{b-a}$ for $(x \geq a) \wedge (x \leq b).$

Expected value:

$$E(\xi) = \int_a^b \frac{u}{b-a}\, du = \frac{1}{2}(a+b).$$

Variance:

$$\mathrm{var}(\xi) = \int_a^b \left(u - \frac{1}{2}(a+b)\right)^2 \frac{1}{b-a}\, du = \frac{1}{12}(b-a)^2.$$



Example:

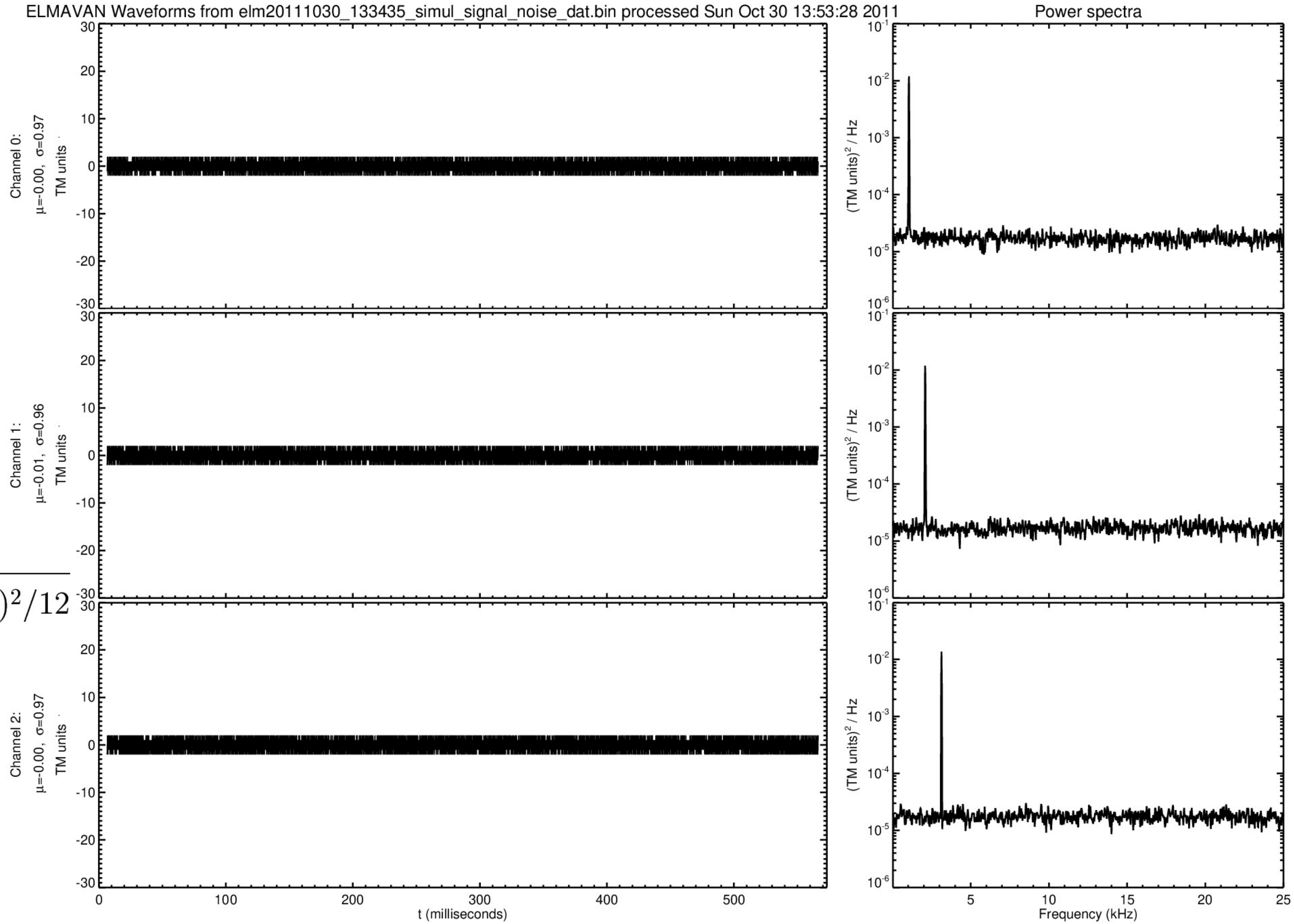
Simulated uniform noise
< -1, +1 >

+

Sine signal with unit amplitude

$$\sigma = \sqrt{\frac{1}{2\pi}\int_0^{2\pi}\sin^2(x)dx + (b-a)^2/12}$$

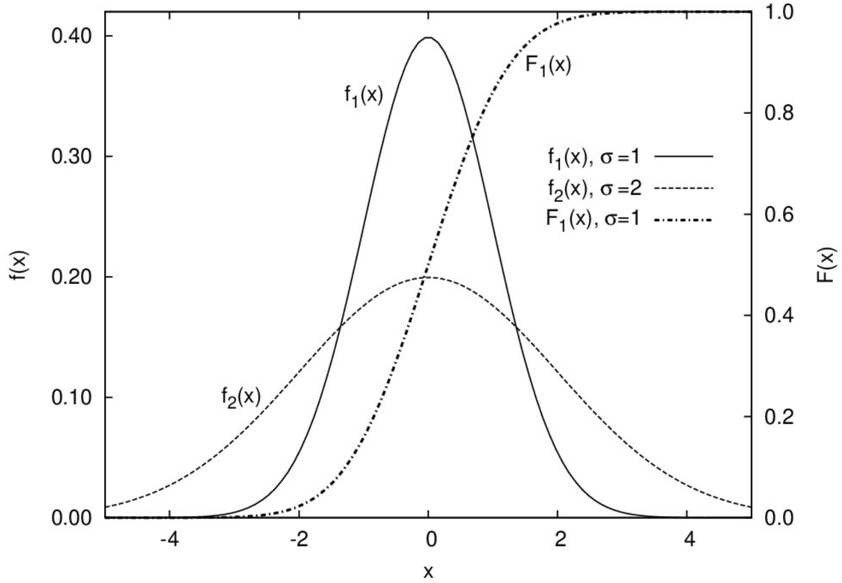
$$\approx \sqrt{1/2 + 2^2/12} \approx 0.9$$



17. Normal distribution

Probability density function:

$$f(u) = n_{\mu,\sigma}(u) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right)$$



Expected value: $E(\xi) = \mu$

Variance: $\text{var}(\xi) = \sigma^2$

Cumulative distribution function:

$$F(x) = N_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \exp\left(-\frac{t^2}{2}\right) dt = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\frac{x-\mu}{\sigma}}{\sqrt{2}}\right)\right]$$

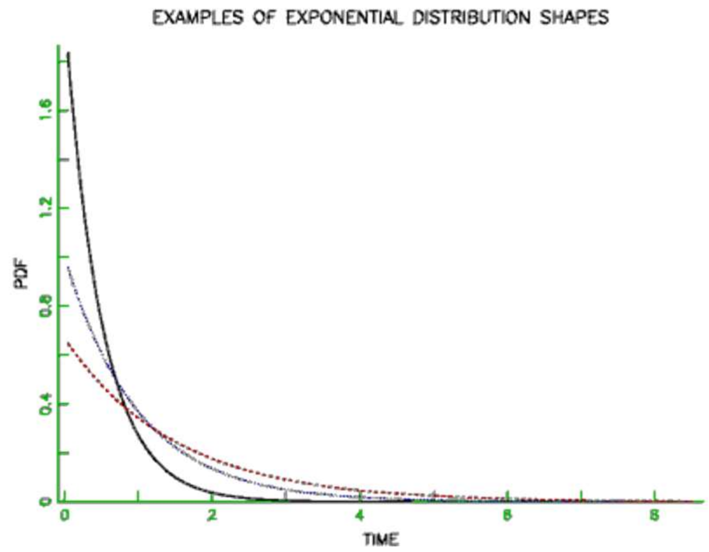
Error function $\text{erf}\,(y) = \frac{2}{\sqrt{\pi}} \int_0^y \exp\left(-t^2\right) dt;$ $\Phi(y) = \sqrt{\frac{2}{\pi}} \int_0^y \exp\left(-\frac{t^2}{2}\right)$ **Laplace function**

Probability within $a\sigma$: $\Phi(a) = P(-a\sigma < \xi - \mu < a\sigma)$:

$\Phi(0.67) = 0.5,$ $\Phi(1) = 0.68,$ $\Phi(2) = 0.95,$ $\Phi(3) = 0.997$

FWHM (full width at half maximum) $= 2\sqrt{2\ln 2}\sigma \approx 2.355\sigma$ $\Phi(\sqrt{2\ln 2}) \approx 0.76$

18. Exponential distribution



Probability density function:

$$f(u) = \frac{1}{\delta} \exp\left(-\frac{u}{\delta}\right) \qquad (\delta > 0,\ u \geq 0).$$

Cumulative distribution function $(x \geq 0)$:

$$F(x) = 1 - \exp\left(-\frac{x}{\delta}\right).$$

Consequence: **distribution without memory**: $P(\xi > a + x \mid \xi > a) = P(\xi > x)$.

Hint for proof: $P(\xi > a + x \mid \xi > a) = P(\xi > a + x) / P(\xi > a) = \big(1 - F(a+x)\big) / \big(1 - F(a)\big)$

Expected value: $E(\xi) = \delta.$

Variance: $\operatorname{var}(\xi) = \delta^2.$

Example: Duration of time intervals between independent events of a Poisson process. (Without memory: If an event has not occurred after 30 seconds, the conditional probability that occurrence will take at least 10 more seconds is equal to the unconditional probability of observing the event more than 10 seconds after the initial time.)

19. Chi-square distribution

Distribution of a random variable $\chi^2 = \sum_{i=1}^{\nu} \frac{(\xi_i - E(\xi_i))^2}{\mathrm{var}(\xi_i)}$, where all ξ_i are normally distributed

ν ... number of degrees of freedom

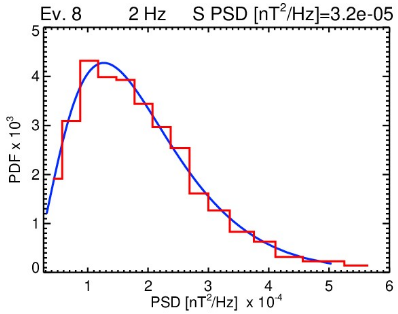
Probability density function:

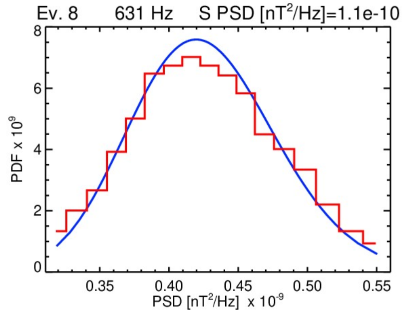
$$f(u) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\, u^{\nu/2-1}\, \exp\left(-\frac{u}{2}\right),$$

where $\Gamma(z) = \int_0^{\infty} t^{z-1} \exp(-t)\, dt.$

Expected value: $E(\xi) = \nu.$

Variance: $\mathrm{var}(\xi) = 2\nu.$





Property: for large ν converges to a normal distribution

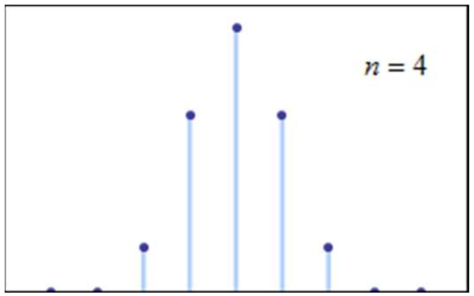
20. Central limit theorem

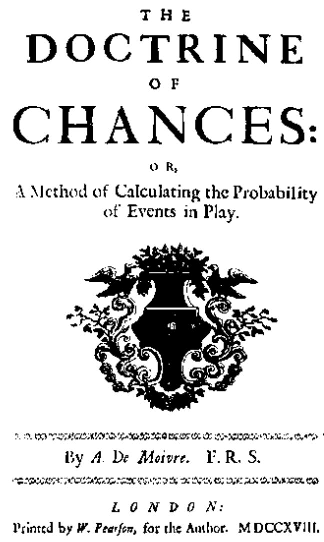
De Moivre–Laplace theorem:

The shape of the discrete binomial distribution based on *n* trials converges to the continuous Gaussian curve of the normal distribution as *n* grows large.

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}, \qquad p+q=1,\ p,q>0$$

for $n \to \infty$




THE DOCTRINE OF CHANCES: OR, A Method of Calculating the Probability of Events in Play.

By A. De Moivre. F. R. S.

LONDON: Printed by W. Pearson, for the Author. MDCCXVIII.

Abraham de Moivre
1667- 1754

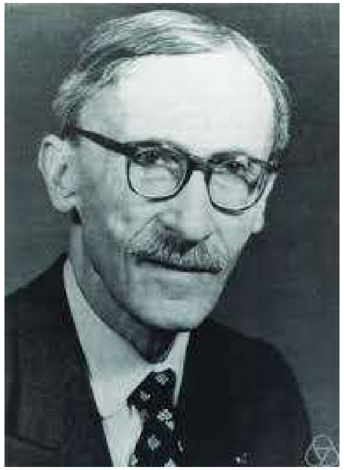
Lindeberg-Lévy central limit theorem:

$\{\xi_n\}$, $n \in \mathbb{N}$ is a sequence of mutually independent random variables with the same probability distributions, with the expected value μ and a finite variance $\sigma^2 < \infty$.

Then, for $x \in \mathbb{R}$

$$\lim_{n\to\infty} P\left\{ \frac{1}{\sigma\sqrt{n}} \left(\sum_1^n \xi_i - n\mu \right) \leq x \right\} = N_{0,1}(x),$$

where $N_{0,1}(x)$ is the cumulative distribution function of a normally distributed random variable with zero expected value and unity variance.



Paul Pierre Lévy
1886-1971

Jarl Waldemar Lindeberg
1876–1932

Ljapunov central limit theorem:

If the following limit for the 3rd central moment is zero: $\lim_{n\to\infty} \frac{\sqrt[3]{\sum E\left(|\xi_i - \mu_i|^3\right)}}{\sqrt{\sum \sigma_i^2}}$ = 0

then, the distributions of the random variables ξ_i do not have to be identical

and the following expression converges to $N_{0,1}(x)$ for $n \to \infty$:

$$\frac{\sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

Aleksandr Mikhailovich Lyapunov
1857- 1918

A simple example of the central limit theorem:

A sum of a large number of random variables with uniform distribution on the interval <0, 1> (each with expected value = 0.5. and variance =1/12) can approximate the normal distribution.

From the Lindeberg – Lévy central limit theorem, a radom variable

$$\sqrt{\frac{12}{n}}\left(\sum_{1}^{n}\xi_i - \frac{n}{2}\right)$$

converges to a normally distributed random variable with zero expected value and unity variance.