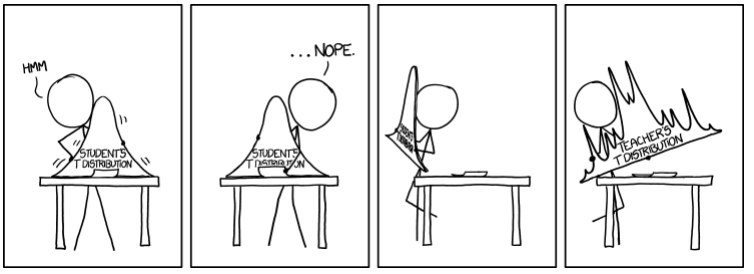
ANALYTICKÁ CHEMOMETRIE

důležitá rozdělení pravděpodobnosti



Patrik Kania
Ústav analytické chemie, VŠCHT Praha

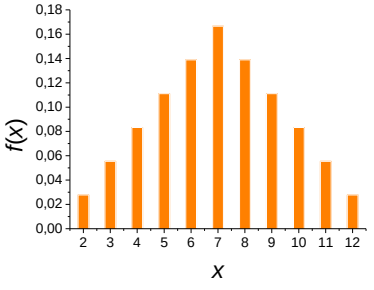
30. 09. 2025

NÁHODNÁ VELIČINA

konečný nebo spočetný počet prvků (např. četnosti)
DISKRÉTNÍ prostor

↓

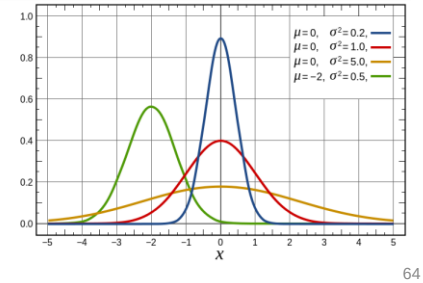
DISKRÉTNÍ rozdělení

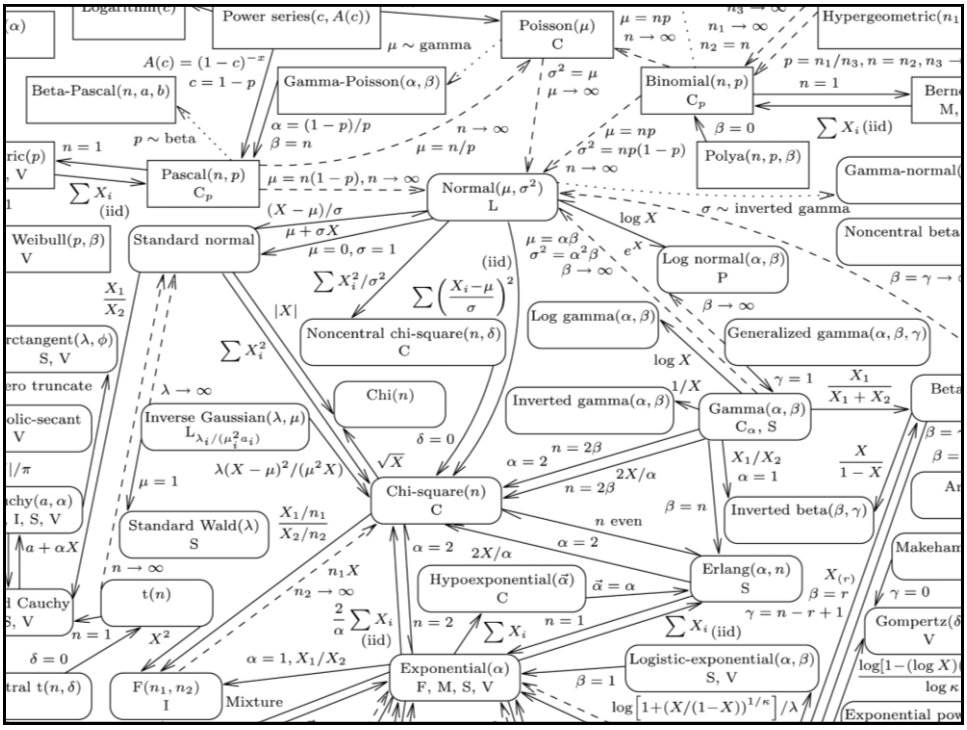


nespočetné množství prvků (většina měřitelných veličin)
SPOJITÝ prostor

↓

SPOJITÉ rozdělení





Diskrétní rozdělení pravděpodobnosti

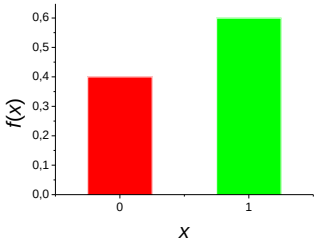
Alternativní rozdělení (angl. *Bernoulli distribution*)

$f(x) = p^x(1-p)^{1-x}, \qquad \text{pro } x = 0 \text{ nebo } 1$

p – pravděpodobnost jevu

tedy $f(1) = p$ a $f(0) = 1 - p$

- náhodná veličina ξ má pouze dva výsledky: elementární jev nastal nebo nenastal (např. hod mincí, úspěch u zkoušky z chemometriky)



$\mathrm{E}(\xi) = p \qquad \mathrm{D}(\xi) = p(1-p)$

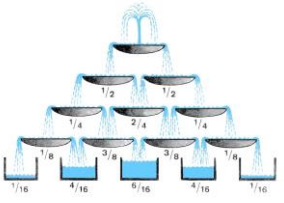
Binomické rozdělení (angl. *binomial distribution*)

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

$$f(x, n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad \text{pro } x = 0, 1, 2, \ldots, n$$

značí se Bi(*n*, *p*) *p* – prst jevu, *n* – počet opakování, *x* – počet výskytů jevu

- opakované alternativní rozdělení
- udává pravděpodobnost *x* výskytů elementárního jevu při *n* opakováních
- pravděpodobnost *p* je při všech opakováních stejná a předem známá
- všech *n* pokusů je vzájemně nezávislých, tj. výsledek předchozího pokusu neovlivňuje výsledky pokusů následujících

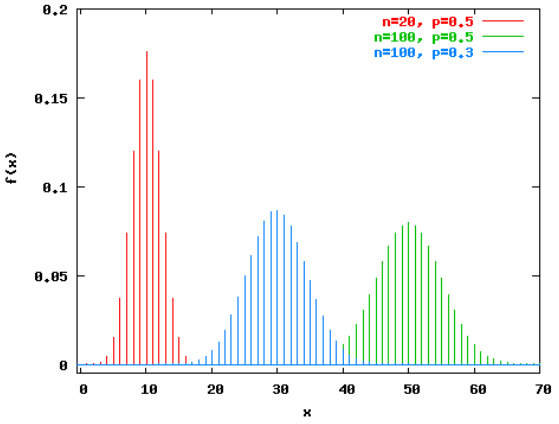


Binomické rozdělení (angl. *binomial distribution*)

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

$$f(x,n,p) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad \text{pro } x = 0, 1, 2, \dots, n$$

značí se Bi(n, p) p – prst jevu, n – počet opakování, x – počet výskytů jevu



$$\mathrm{E}(\xi) = np$$

$$\mathrm{D}(\xi) = np(1-p)$$

pro $n = 100$
$p = 0{,}3$

je $\mathrm{E}(\xi) = 30$
$\mathrm{D}(\xi) = 21$

Vztahy mezi uvedenými diskrétními rozděleními:

alternativní

↓ opakování pokusu, tedy $n > 1$

binomické

↓ pro $n \to \infty$

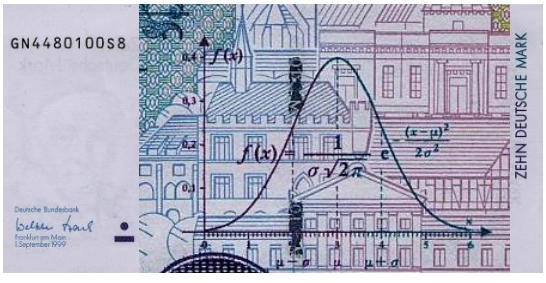
normální SPOJITÉ!

Spojitá rozdělení pravděpodobnosti

Normální rozdělení N(μ, σ^2) (angl. *normal distribution*)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{pro } -\infty < \mu < \infty, \sigma^2 > 0$$

- mnoho náhodných veličin se řídí přesně tímto rozdělením
- za určitých podmínek ale dobře aproximuje i jiná rozdělení
- dobrá aproximace chyb měření → zákon chyb

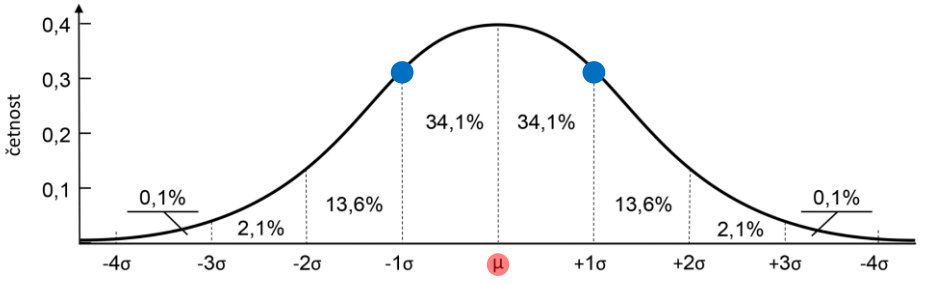


má dva parametry:

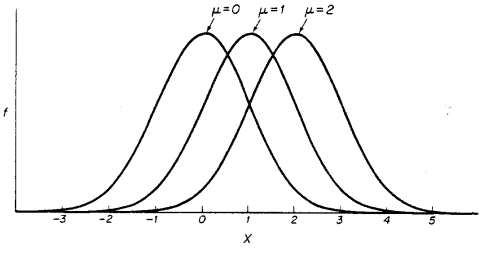
střední hodnotu μ
rozptyl σ^2

základní vlastnosti:

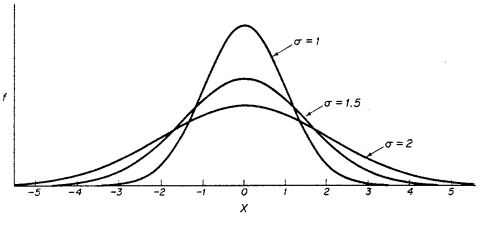
- je souměrné dle osy $x = \mu$ a v bodě $x = \mu$ má maximum
- má inflexní body v bodech $\mu + \sigma$ a $\mu - \sigma$
- střední hodnota je i mediánem a modem
- nekonečné množství kombinací μ a σ



vliv změny střední hodnoty



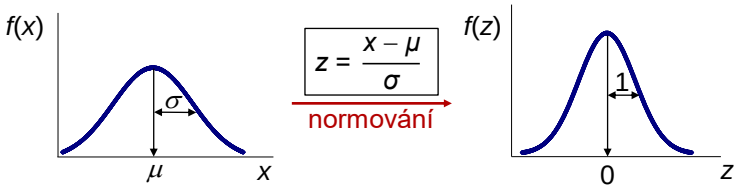
vliv změny rozptylu



standardizace rozdělení

náhodná veličina ξ
normální rozdělení $N(\mu, \sigma^2)$

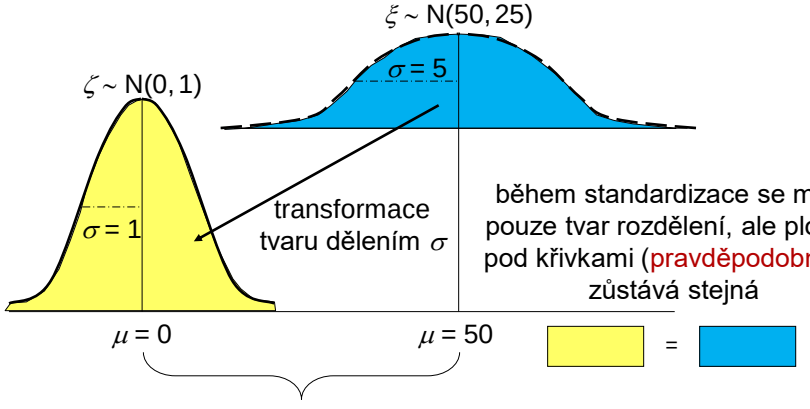
náhodná veličina ζ
normované normální rozdělení $N(0, 1)$



distribuční funkce $F(x) = P\{\xi \leq x\}$

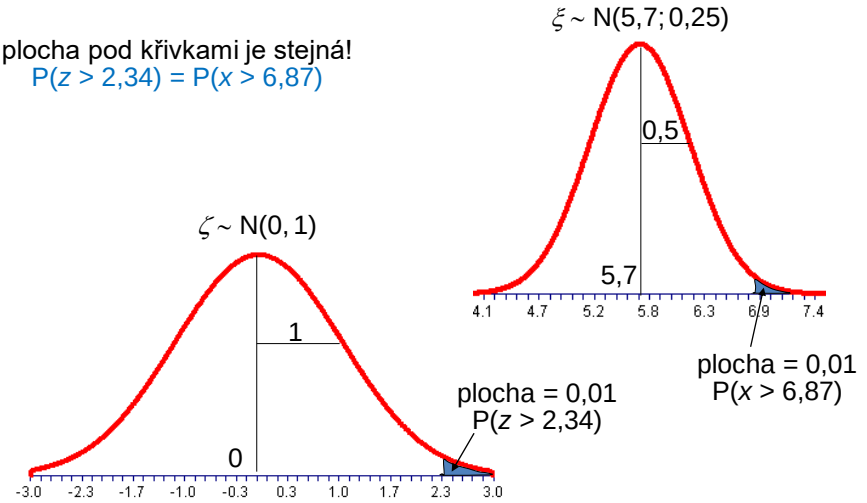
distribuční funkce $\Phi(z) = P\{\zeta \leq z\}$

standardizace rozdělení



standardizace rozdělení

plocha pod křivkami je stejná!
P(z > 2,34) = P(x > 6,87)



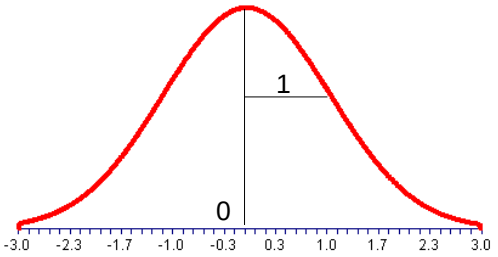
Distribuční funkce normálního rozdělení N(0, 1) Excel: **NORMSDIST(*z*)**
(pro hodnoty *z* jsou uvedeny příslušné hodnoty F(*z*))

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78527
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84850	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92786	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99001	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99586	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99830	0,99836	0,99841	0,99846	0,99851	0,99856	0,99860
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983

Normované normální rozdělení N(0, 1) (angl. *standard normal distrib.*)

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad \text{pro } -\infty < z < \infty$$

- jednotné rozdělení pro všechno normální
- snazší výpočet hustoty pravděpodobnosti a distribuční funkce
- lze tabelovat → rychlý výpočet úlohy



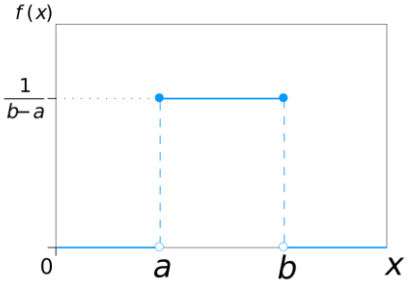
má dva konstantní parametry:

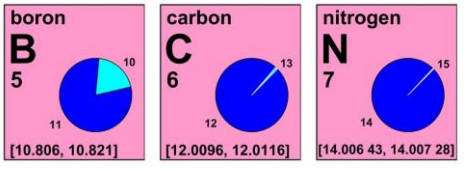
střední hodnotu 0
rozptyl 1

Rovnoměrné rozdělení (angl. *uniform distribution*)

$$f(x) = \frac{1}{b-a}, \qquad \text{pro } a \le x \le b$$

- všechny hodnoty náhodné veličiny ξ jsou stejně pravděpodobné (např. rozdělení hmotností celé periodické tabulky)





Spojitá rozdělení pravděpodobnosti

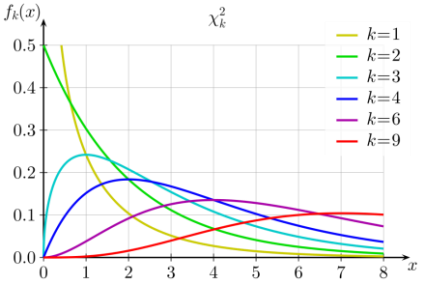
- uměle vytvořená -

χ^2 - rozdělení (chí-kvadrát či Pearsonovo) (angl. *chi-squared distribution*)

$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} e^{(-\frac{x}{2})} x^{(\frac{k}{2}-1)}, \qquad \text{pro } x > 0$$

k – počet stupňů volnosti

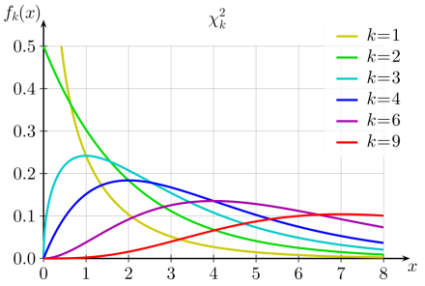
Jsou-li $X_1, \dots, X_k$ nezávislé náhodné veličiny s rozdělením N(0, 1), pak veličina $Q = X_1^2 + X_2^2 + \cdots + X_k^2$ má rozdělení $\chi^2(k)$



$$\mathrm{E}(\xi) = k \qquad \mathrm{D}(\xi) = 2k$$

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} \underbrace{\frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}}_{K} \underbrace{e^{(-\frac{x}{2})}}_{1} x^{(\frac{k}{2}-1)} = K \lim_{x\to 0+} x^{(\frac{k}{2}-1)}$$

- rozdělení je nesouměrné



pro $k = 1 \rightarrow K \lim_{x\to 0+} \frac{1}{\sqrt{x}} = \infty$

$k = 2 \rightarrow K \lim_{x\to 0+} x^0 = K$

$k = 3 \rightarrow K \lim_{x\to 0+} \sqrt{x} = 0$

Mějme normální náhodnou veličinu ξ s rozdělením N(μ, σ^2) a proveďme výběr prvků/hodnot této veličiny v rozsahu *k*. Pro každý získaný výběr vypočítejme hodnotu Y_j

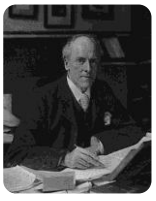
$$\boxed{Y_j = \sum_{i=1}^{k} \left(\frac{x_i - \mu}{\sigma}\right)^2 = \sum_{i=1}^{k} z_i^2}$$

- získané hodnoty Y_j jsou hodnotami náhodné veličiny s $\chi^2(k)$ rozdělením
- je to rozdělení sumy čtverců normované náhodné veličiny
- počet vybraných prvků *k* je roven počtu stupňů volnosti
- hustota pravděpodobnosti není pro nás zajímavá
- zajímavé jsou pouze kritické hodnoty $\chi^2_\alpha(k)$ či kvantily $\chi^2_{1-\alpha}(k)$

Friedrich Robert Helmert (1843–1917)

- německý geodet a statistik
- v roce 1876 objevuje rozdělení výběrového rozptylu normálního rozdělení
- publikace v němčině → pro Anglii neznámý

Helmertovo rozdělení vs. chí-kvadrát rozdělení



Karl Pearson (1857–1936)

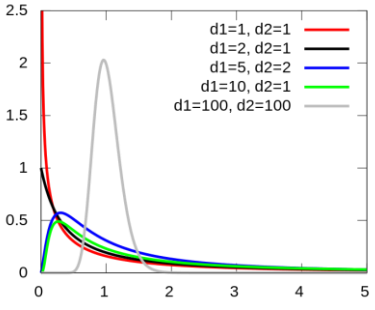
- anglický matematik a filozof
- v roce 1911 založil v Londýně první katedru statistiky
- v roce 1900 objevuje (nezávisle) stejné rozdělení
- publikace v angličtině → světoznámý

F-rozdělení (Fisher-Snedecorovo) (angl. *F-distribution*)

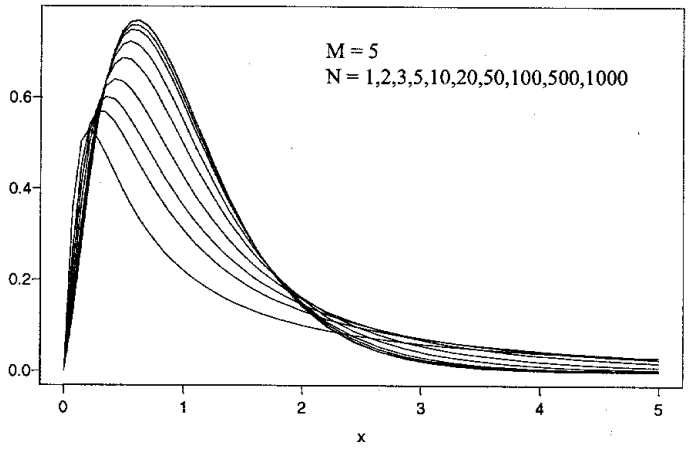
Jsou-li X a Y nezávislé náhodné veličiny s rozdělením $\chi^2(m)$ a $\chi^2(n)$ a stupni volnosti m a n, pak náhodná veličina F

$$F = \frac{\frac{X}{m}}{\frac{Y}{n}}$$

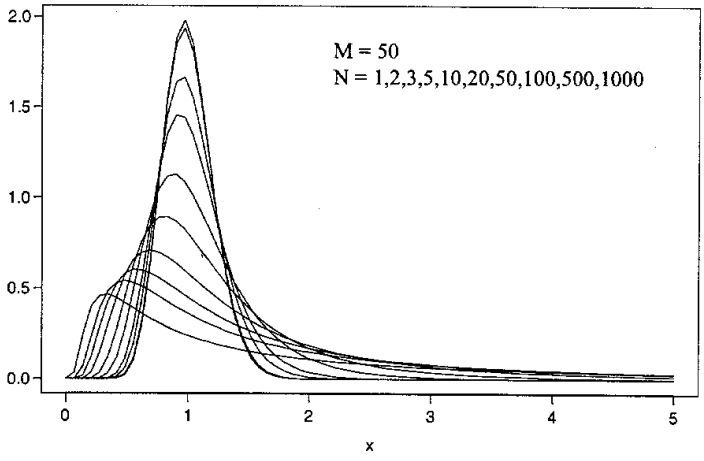
má právě F-rozdělení F(m, n)



Hustota pravděpodobnosti F-rozdělení F(M, N)



Hustota pravděpodobnosti F-rozdělení F(M, N)



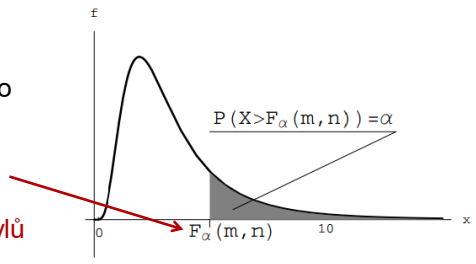
F-rozdělení (Fisher-Snedecorovo) (angl. *F-distribution*)

Jsou-li *X* a *Y* nezávislé náhodné veličiny s rozdělením $\chi^2(m)$ a $\chi^2(n)$ a stupni volnosti *m* a *n*, pak náhodná veličina *F*

$$F = \frac{\frac{X}{m}}{\frac{Y}{n}}$$

má právě F-rozdělení F(*m*, *n*)

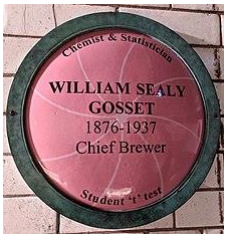
- hustota pravděpodobnosti není pro nás zajímavá
- zajímavé jsou kritické hodnoty $F_\alpha(m,n)$ či kvantily $F_{1-\alpha}(m,n)$
- používá se pro test rovnosti rozptylů

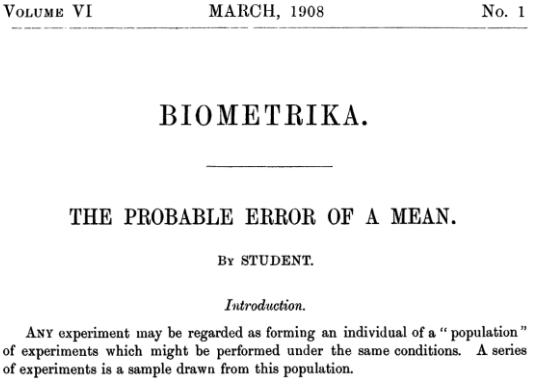


POZOR! $F_\alpha(m,n) \neq F_\alpha(n,m)$

William Sealy Gosset (1876–1937)

- anglický statistik, chemik a sládek
- zájem o analýzu průměrů na malých vzorcích
- v roce 1906–1907 stáž u Pearsona
- roku 1912 je Studentovo rozdělení korigováno Fischerem





VOLUME VI MARCH, 1908 No. 1

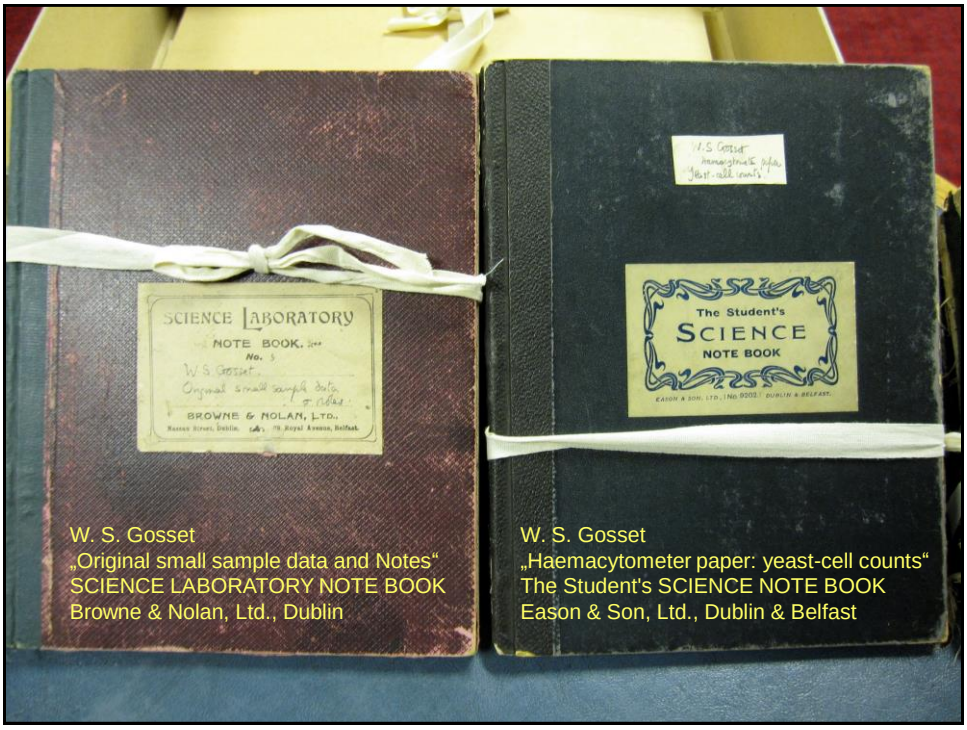
BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

Introduction.

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.



W. S. Gosset
„Original small sample data and Notes“
SCIENCE LABORATORY NOTE BOOK
Browne & Nolan, Ltd., Dublin

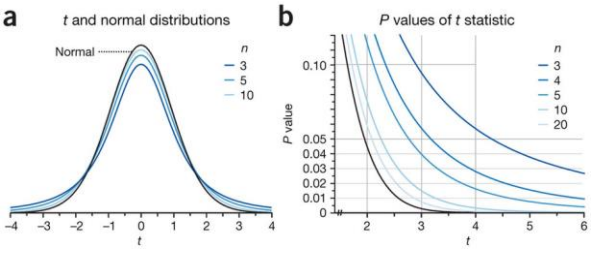
W. S. Gosset
„Haemacytometer paper: yeast-cell counts“
The Student's SCIENCE NOTE BOOK
Eason & Son, Ltd., Dublin & Belfast

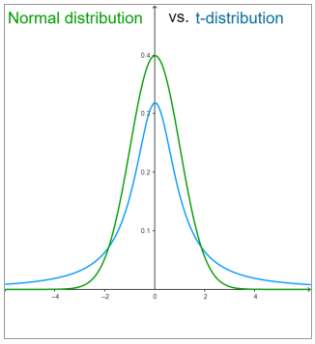
t-rozdělení (Studentovo) (angl. *Student's t-distribution*)

Jsou-li X a Y nezávislé náhodné veličiny, přičemž X má rozdělení $N(0,1)$ a Y má rozdělení $\chi^2(k)$, pak náhodná veličina T

$$T = \frac{X}{\sqrt{\frac{Y}{k}}}$$

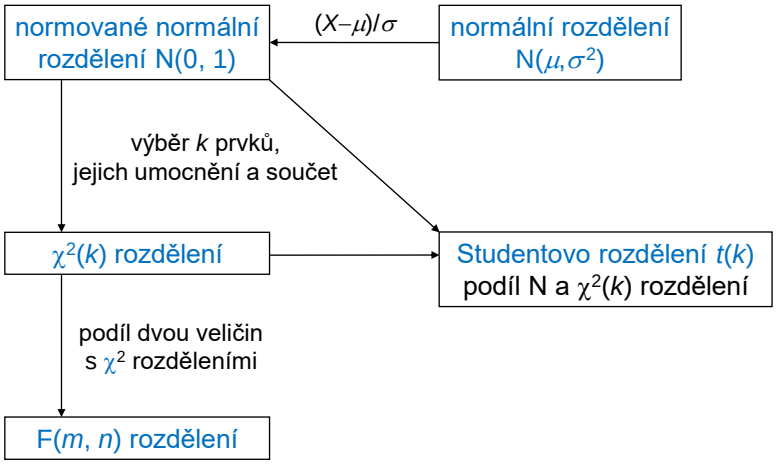
má Studentovo rozdělení $t(k)$ s k stupni volnosti





- je rozdělení výběrového průměru
- počet stupňů volnosti $n - 1$
- pro $n \to \infty$ (v praxi pro $n > 30$) přechází v normální rozdělení $N(0, 1)$
- použití tam, kde neznáme σ^2 náhodné veličiny a nahrazujeme ho s^2
- tabelace kritických hodnot $t_\alpha(n)$ či kvantilů $t_{1-\alpha}(n)$

Vztahy mezi uvedenými spojitými rozděleními:



Konec