



The International Statistical Institute (ISI) is one of the oldest international scientific associations functioning in the modern world. Its first congresses were convened in 1853, and it was formally established in 1885. The Institute is an autonomous society, which seeks to develop and improve statistical methods and their application through the promotion of international activity and co-operation.

### **ISI Sections:**

- [Bernoulli Society](#)
- [International Association of Survey Statisticians](#)
- [International Association for Statistical Computing](#)
- [International Association for Official Statistics](#)
- [International Association for Statistical Education](#)
- [Irving Fisher Society for Financial and Monetary Statistics \(ISI transitional Section\)](#)

**Stephen M. Stigler**  
**President of ISI, Professor of Statistics,**  
**University of Chicago:**



„Statisticians agree on the importance of statistics for national prosperity, but what is the basis of this belief? There does seem to be a strong historical association between statistics and prosperity... In the 17th that century in France, Blaise Pascal and Pierre de Fermat were exploring an early probability theory that foreshadowed modern financial engineering; while in Holland, Christiaan Huygens worked on many of the same problems and Jan de Witt was taking first steps toward a mathematical framework for life insurance. To a degree, the pursuit of all of these statistical sciences in these European nations rose and fell with their economic fortunes. With the rise of British sea trade and the decline of Dutch trade, there was a rise in British statistics while Dutch work in that area languished. As French influence increased in the late 1700's and with the Napoleonic era, so too did their statistics, both mathematical (Laplace) and national (from Turgot and Necker to the development and spread of statistical accounting under Napoleon and even afterwards, in the compilations of Fourier and Chabrol). So too, the unification and growth in economic power of Germany was accompanied by a great increase in the collection and analysis of economic statistics, from the early national statistics around 1800 (Lueder, Meusel) to the vast Prussian publications towards the end of the nineteenth century. The formal introduction of regular censuses also was tied to national growth, in the United States, Great Britain, and France, where censuses began to take hold as growing territories – empires in some cases – came to cope with the need to “number the people” for all sorts of reasons, from taxation to conscription to determining necessary social services. Smaller nations and states exhibited the same association, though on a smaller scale. Conversely, one might cite the example of the general weakness of Soviet statistics leading up

to the 1980's, notwithstanding their extraordinary strength in mathematics and theoretical probability.“

## **Jak se historicky vyvinulo statistické myšlení:**

Tři vědecké problémy 18. století, úzce spojené s metodou nejmenších čtverců:

- (1) **určit a matematicky popsat pohyby měsíce,**
- (2) **vysvětlit neperiodické odchylky v pohybu planet Jupiter a Saturn**
- (3) **popsat přesný tvar zeměkoule.**

Idea: vysvětlit tyto jevy pomocí Newtonovy gravitační teorie. Studium a objasnění těchto problémů je mj. spojeno se slavnými jmény **Mayer, Euler, Legendre, Gauss a Laplace.**

## **Johann Tobias Mayer**



*(1723-1762, narozen v Göttingenu)*

**Tobias Mayer** byl matematik-samouk, který pracoval jako kartograf v Norimberku. Proslavil se tím, že odhalil kývavý pohyb (the

libration) měsíčního disku, a díky tomu byl v r. 1751 jmenován profesorem ekonomie a matematiky v Göttingenu. V letech 1748-49 provedl mnoho pozorování měsíce a popsal jeho pohyby 27 nekonsistentními rovnicemi o třech neznámých. Jako praktik byl přesvědčen, že přesnost výsledku lze zvýšit vhodnou kombinací jednotlivých pozorování, a navrhl statistické řešení problému, jak vhodně pozorování (a rovnice) kombinovat a nakonec je reprezentovat třemi rovnicemi o třech neznámých.

## Leonhard Euler



*(1707-1783, narozen v Basileji)*

Naproti tomu **Leonard Euler** byl hlavně vynikajícím matematikem. Je autorem knih o variačním počtu, o výpočtu drah planet, o střelbě a balistice, o stavbě lodí a navigaci, o pohybu měsíce. V r. 1748, kdy působil v Berlíně, pařížská Akademie věd vypsala cenu na vysvětlení nepravidelností ve vzájemném pohybu Saturnu a Jupitera. Euler provedl matematickou analýzu pohybu těchto planet po eliptických drahách. Svou analýzu chtěl ověřit na empirických datech, čímž získal 75 rovnic, nejprve s 8 neznámými. Jako exaktní

matematik však nepřijal podobnou myšlenku jako Tobias Mayer, vzájemně kombinovat rovnice a pozorovaná data, a tím získat lepší popis a výklad jevů.

## Pierre-Simon Laplace



**(1749-1827, narozen v Beaumont-en-Auge, Normandie)**

Ve své analýze pohybu Saturnu a Jupitera Laplace navázal na Eulera a kombinací nekonsistentních lineárních rovnic i na Mayerovu metodu, a odhalil periodicitu pohybů. Zatímco Mayer pouze sčítal rovnice uvnitř určitých skupin, Laplace vzájemně kombinoval tytéž rovnice různými způsoby a tím se dostal blíže k obecnému statistickému řešení problému. Své závěry publikoval v r. 1787 ve 127-stránkovém memoáru „***Théorie de Jupiter et de Saturne***“. Matematická astronomie a teorie pravděpodobnosti byly hlavní náplní jeho práce během jeho života.

## Adrien-Marie Legendre



*(1752-1833, narozen v Paříži)*

V r. 1806 Legendre publikoval knihu o určování dráhy komet: *Nouvelles méthodes pour la détermination des orbites des comètes*, ve které napsal: „Uvažoval jsem, že v problému komet je lépe začít od okamžitě napozorovaných dat, a všemi možnými prostředky co nejvíce zjednodušit formule a rovnice, které vedou k určení jednotlivých prvků dráhy komety.“ Legendre předpokládal, že kometa sleduje parabolickou dráhu. Svou metodu aplikoval na známá data o dvou kometách. Protože sledoval po třech pozorováních ve stejně dlouhých intervalech, získal soustavu rovnic, která měla více rovnic než proměnných. Nesnažil se ji řešit jako soustavu, ale v appendixu ke své knize Legendre navrhl metodu nejmenších čtverců, která vyrovnávala napozorovaná data s předpokládanou křivkou. Shodou okolností, Gauss publikoval svou verzi metody nejmenších čtverců v r.1809; a ačkoliv se zmiňoval, že metoda byla popsána v Legendreově knize, Gauss připisoval prioritu sobě. Legendre se tím cítil hluboce poraněn a



dlouho bojoval za uznání své priority. 10 let po publikaci Legendreova appendixu se metoda nejmenších čtverců stala standardní metodou astronomie a geodézie ve Francii, Itálii a Prusku.

## Rudjer Josef Boskovič



Z jugoslávské známky



Chorvatská bankovka

**(1711-1787, narozen v Raguse, nyní Dubrovník)**

Boskovič studoval na Collegium Romanum v Římě and v r. 1740 zde byl jmenován profesorem matematiky. Napsal 70 článků o optice, astronomii, gravitaci, meteorologii a trigonometrii. Tvar zeměkoule byl historicky studován dvěma metodami: pozorováním pohybu kyvadla na různých místech planety a měřením délky oblouku stejného úhlu na témže poledníku ale na různých, vzájemně vzdálených místech. V r. 1735 Francouzská akademie uspořádala expedice do Peru a do Laplandu, které měly porovnat úhly blízko rovníku a na  $66^\circ$  s měřením v okolí Paříže. V r. 1755 **Boskovič** spolu s anglickým jesuitou **Christopherem Mairem** publikoval výsledky měření poledníkového úhlu v okolí Říma pod názvem ***Letteraria Expeditione per Pontificiam ditionem ad dimetiendas duas Meridiani gradus.*** Boskovičova analýza těchto dat je prvním úspěšným statistickým postupem a prvním závěrem získaný z nekonsistentních měření. Jeho myšlenkou bylo minimalizovat vážený součet absolutních hodnot odchylek měření

od hledané hodnoty; vážený proto, že jednotlivá měření se lišila svou přesností. Tím vlastně dal základ tzv. **robustním statistickým metodám**, které se plně rozvinuly až ve 2. polovině 20. století.

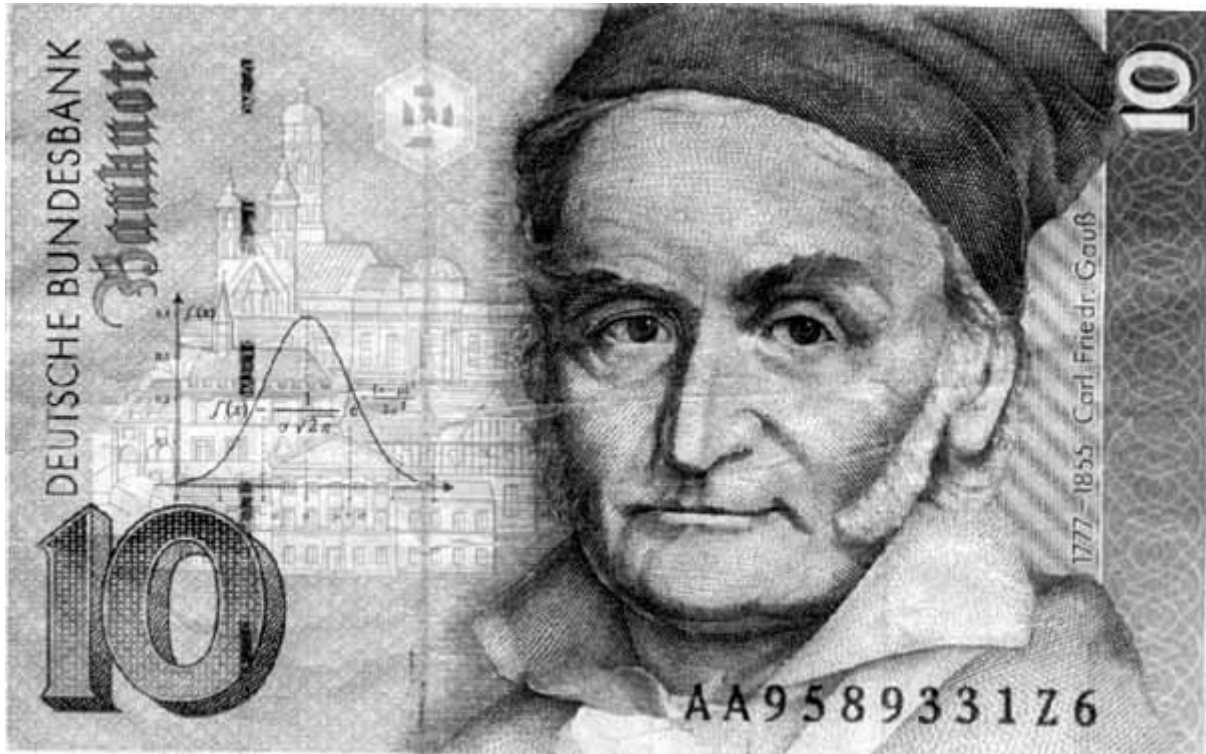
	Místo	Zeměpisná šířka	Délka oblouku (toises)
1	Quito	0°0′	56,751
2	Mys dobré naděje	33°18′	57,037
3	Řím	42°59′	56,979
4	Paříž	49°23′	57,074
5	Lapland	66°19′	57,422

(1 toise ≈ 6.39 stopy)

Na Boskoviče navázal **Laplace**: druhý díl jeho knihy **Mécanique Céleste** obsahuje studii tvaru zeměkoule včetně diskuse dat získaných z několika expedicí, na jejichž analýzu Laplace aplikoval svou teorii chyb pozorování. Pozdější vydání jeho knihy **Théorie Analytique des Probabilités** též pokrývá aplikace pravděpodobnosti na určení hmotnosti Jupitera, Saturnu a Uranu a na problémy geodésie se speciálním zaměřením na měření francouzského poledníku.



## Johann Carl Friedrich Gauss



*(1777-1855, narozen v Brunswicku, Německo)*

Geniální matematik Gauss uvažoval podobné problémy jako Mayer, Euler, Boskovič, Laplace an Legendre, avšak rigorózně z hlediska teorie preděpodobnosti. Hledal rozdělení pravděpodobností chyb měření, pro které je odhad metodou nejmenších čtverců (nebo průměr ve speciálním případě) optimálním odhadem měřené fyzikální (nebo jiné) konstanty, a dospěl ke křivce (hustotě) rozdělení pravděpodobností tvaru

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

Toto rozdělení se dnes nazývá **normální** nebo **Gaussovo**. Stephen Stigler v tom spatřuje další uplatnění zákona „**Stigler's Law of Eponymy**”, který říká, že „**Žádný vědecký objev není**

**pojmenován po svém objeviteli“.** Tuto křivku chyb zmiňuje Laplace již v r. 1774, a ve skutečnosti poprvé ji popsal v r. 1733 **Abraham de Moivre**, jehož dílo znali Gauss i Laplace. Toto rozdělení pravděpodobností hrálo centrální roli v matematické statistice od r. 1810, a v praxi se předpokládalo, že většina náhodných veličin se jím řídí, čímž vzniklo „dogma normality“.

Výrok přisuzovaný francouzskému fyzikovi G. Lippmanovi (H. Poincaré, *Calcul des Probabilités*, Paříž 1912): „Všichni věří v normální rozdělení chyb: experimentátoři, protože je pokládají za matematický teorém, a matematikové, protože je pokládají za experimentální fakt.“

Toto dogma bylo potlačeno až v době výkonných počítačů, přestože již v r. 1965 **Kagan, Linnik** (SSSR) a **C.R. Rao** (Indie) dokázali a publikovali v indickém časopise *Sankhya* výsledek, že odhad metodou nejmenších čtverců je optimální pouze pro normální rozdělení chyb, zatímco v ostatních případech může úplně selhat. Legendre ve své původní publikaci o metodě nejmenších čtverců z r. 1805 doporučoval nejprve „odmítnout měření, která jsou příliš velká, než aby mohla být pokládána za přípustná“.

Syntéza Gaussových a Laplaceových výsledků poskytla účinný nástroj pro astronomii a geodézii. Připomeňme např. stanovení střední hustoty Země a gravitační konstanty [H. Cavendish: „*Experiments to determine the density of the earth*“ (Philosophical Trans. of the Royal Society of London for the year 1798), Laplace: „*Sur la densité moyenne de la terre*“ (Ann. Chemie et de physique 14, 1820)] nebo měření rychlosti světla ve 2 polovině 19. století (Foucault, Michelson a Newcomb).

## Simon Newcomb



**(1835-1909, narozen ve Wallace, Nova Scotia, Canada)**

Byl profesorem matematiky a astronomie v Johns Hopkins University od 1884 do 1893, editorem časopisu *American Journal of Mathematics* a zakládajícím členem a prvním prezidentem American Astronomical Society; byl též prezidentem American Mathematical Society. Byl i nadaným ekonomem.

Newcomb shromáždil soubor 684 pozorování tranzitu Merkura a zjistil, že rozdělení pravděpodobností jejich reziduí je mnohem “těžší” než normální, a navrhl modelovat data pomocí směsi dvou normálních rozdělení s různými parametry; tím se stal i průkopníkem robustních statistických metod.

Ve společenských vědách však tyto metody byly začátkem 19. století zcela neznámé a pro tuto oblast byly objeveny mnohem později.

## Zápas o uplatnění teorie pravděpodobnosti ve společenských vědách

Lambert Adolphe Jacques Quetelet



(1796-1874, narozen v Gentu, Flandry)

Vzděláním původně matematik, později doplnil své vzdělání v astronomii a v meteorologii a vedl Královskou observatoř v Bruselu. Pravděpodobnost a její aplikace se naučil od Josepha Fouriera a od samotného Laplace, a mezinárodní reputace dosáhl hlavně jako statistik a sociolog. Od Laplace převzal metodu sledování změn v populaci, použitou ve Francii, založenou nikoli na úplném sčítání lidu, ale na sčítání v souboru náhodně vybraných oblastí země nebo populace. V r. 1835 shrnul své poznatky v knize „*Sur l'homme et le développement de ses facultés, ou essai de physique sociale*“, kde Quetelet prezentoval svůj pojem *průměrného člověka* jakožto centrální hodnoty, kolem které se měřené lidské vlastnosti shlukují podle normální křivky. V r. 1853

Quetelet organizoval prvú mezinárodnú štatistickú konferenciu. Mezinárodne užívaná miera obezity je **Queteletův index**, někdy nazývaný *Body mass index* (BMI), definovaný ako  $QI = (\text{váha v kilogramoch})/(\text{výška v metroch})^2$ . Jestliže  $QI > 30$ , je osoba oficiálne obezná. Zajímavé sú jeho štúdie kriminality a pomeru usvedčených zločinců.

## Ladislaus Josephowitsch Bortkiewicz



**(Narodený 1868 v St Petersburgu, zomrel 1931 v Berlíne)**

Von Bortkiewicz študoval právo v St Petersburgu, a následne politickú ekonomiu a štatistiku. Doktorát získal v Göttingenu. Jeho práce v aktuárskych vedách sa týkala hlavne úmrtnostných tabuliek. V r. 1898 publikoval prácu „**The Law of Small Numbers**“, venovanou predovšetkým Poissonovu rozdeleniu pravdepodobností, ktoré snad malo byť nazvané jeho menom. Toto dielo je snad najviac známe vzhľadom k príkladu, v ktorom ukázal, že počet pruských dôstojníkov, ktorí v období 1875-1894 zomreli na následky kopnutí konňami, sa dobre riadi Poissonovým rozdelením pravdepodobností. Von Bortkiewicz si uvedomil, že javy s malou frekvenciou sa v veľkej populácii riadi Poissonovým zákonom, i keď sa pravdepodobnosti jednotlivých javů líšia.

## Vilfredo Pareto



*Vilfredo Pareto*

*(1848-1923, narozen v Paříži)*

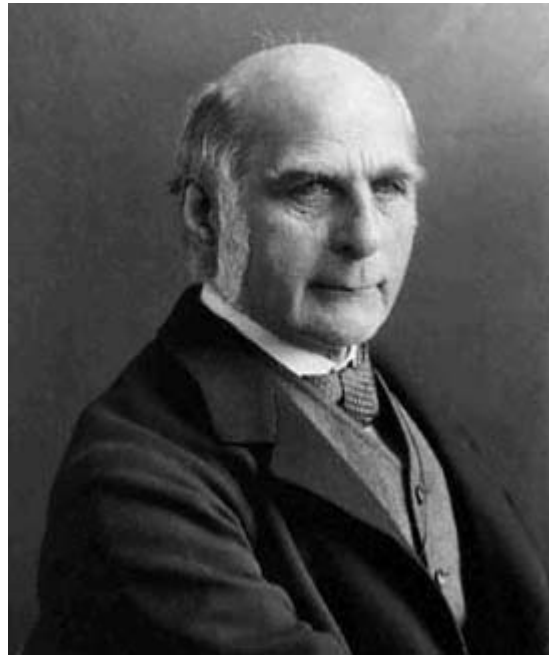
Italský ekonom Vilfredo Pareto studoval inženýrství na Polytechnickém Institutu v Turinu. Jeho hlavním ekonomickým i statistickým příspěvkem je "Paretův zákon" rozdělení příjmů, založený na argumentaci, že ve všech zemích a ve všech časových epochách se rozdělení příjmu a blahobytu řídí logaritmickým modelem podle formule  $\log N = \log A + m \log x$ , kde  $N$  je počet s příjmem vyšším než  $x$ , a  $A$  a  $m$  jsou konstanty modelu. Během doby byl Paretův zákon mnohokrát empiricky potvrzen, a dodnes je základním modelem v pojišťovnictví i jinde.



# BIOMETRICKÁ ŠKOLA

## Studium dědičnosti

### Francis Galton



*(1822-1911, narozen ve Sparkbrooku blízko Birminghamu)*

Francis Galton je znám svými pionýrskými studii lidské inteligence. V r. 1840 začal studovat medicínu na Trinity College, Cambridge, ale brzy přešel na matematiku. Galton byl bratrancem Charlese Darwina a velice ho ovlivnila jeho publikace *O původu druhů* (1859). Založil laboratoř pro získávání statistických dat na lidských jedincích. Na rozdíl od sociální statistiky mohl své experimenty vhodně naplánovat. Shromáždil řadu dat o výšce, váze a síle lidí. Laboratoř byla předchůdkyní Biometrické laboratoře Karla Pearsona na University College, London. Galton uspořádal grafy výšek rodičů a výšek jejich dospělých dětí a poprvé sestrojil tzv. regresní přímku. V r. 1888 studoval velikost různých orgánů těžce osoby a poprvé definoval korelační index dvou vlastností. Své výsledky shrnul v r. 1889 v publikaci „**Natural inheritance**“. Ve své laboratoři shromáždil také údaje o otiscích prstů a ukázal, že tyto se s věkem nemění a mohou být použity k identifikaci (viz jeho publikace **Finger prints** (1893), **Blurred finger prints** (1893), and

***Finger print directory*** (1895)). Jeho identifikační systém se stal základem klasifikace Sira Edwarda R. Henryho, později hlavního komisaře londýnské metropolitní policie. Významné jsou též Galtonovy příspěvky k meteorologii, antropometrii a fyzikální antropologii. V r. 1863 publikoval stať ***Meteorographica, or methods of mapping the weather***. Zavedl termín anticyklon a ukázal jeho význam v předpovídání počasí.

## Francis Ysidro Edgeworth



**(1845-1926, narozen v Edgeworthstownu, Irsko, zemřel v Oxfordu)**

Studoval jazyky, logiku a patrně i právo a byl profesorem ekonomie. Matematiku a statistiku Edgeworth studoval jako samouk. V r. 1885 publikoval ***Methods of Statistics***, kde ilustroval aplikace a interpretace statistických testů významnosti shodnosti dvou průměrů. Navázal na Galtona a využil jeho výsledky ve společenských vědách. Edgeworth (1887): „Metoda nejmenších čtverců je patrně náš nejlepší kurz, pokud jsme hodili přes palubu určitý podíl našich dat – oběť, kterou často musejí přinést ti, kteří

plují v bouřlivém moři pravděpodobnosti“ („*The choice of means*“ Philosophical Magazine 24).

## Karl Pearson



(1857-1936, narozen v Londýně)

V r. 1879 vystudoval matematiku v Cambridge, pak studoval fyziku a metafyziku v Heidelbergu. V Berlíně sledoval přednášky o darwinismu fyziologa Emil du Bois-Reymonda. Galtonova kniha jej přivedla k zájmu o konstrukci nových matematických metod vhodných ke studiu dědičnosti a evoluce. V letech 1893 až 1912 napsal 18 článků nazvaných ***Mathematical Contributions to the Theory of Evolution***, které obsahují jeho nejvýznamnější výsledky o regresní analýze, koeficientu korelace a obsahují známý Pearsonův chi-kvadrát test (1900). Pearson byl spoluzakladatelem, Spolu s Weldonem a Galtonem založil statistický časopis ***Biometrika***, jehož editorem byl 35 let až do své smrti. Biometrika vychází dodnes.

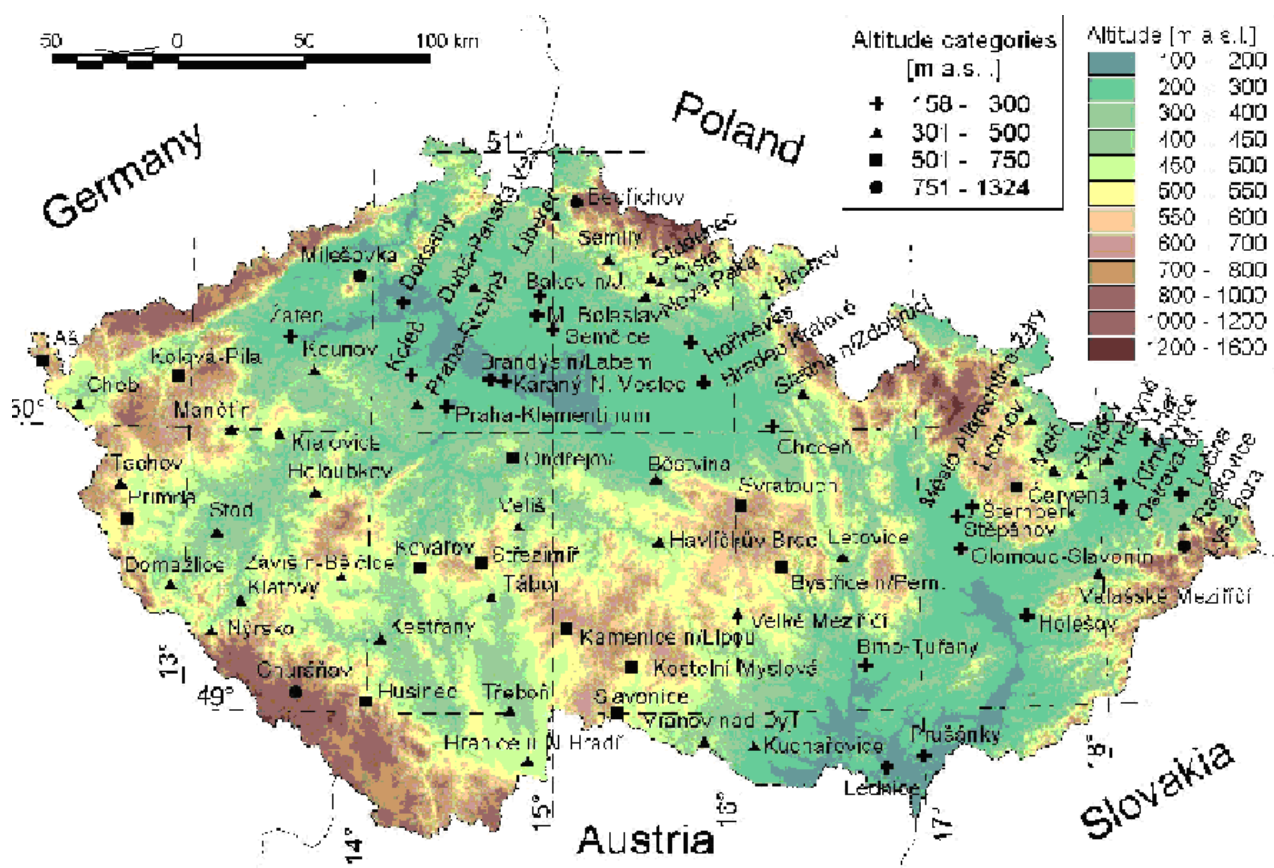
**20. století: Egon S. Pearson, R.A. Fisher, a mnoho dalších...**

## Kde se používá statistika dnes:

Uvedme jen oblasti aplikací, na kterých se naše skupina podílela v poslední době:

Tradičně v medicíně, v meteorologii, ve studiu životního prostředí, velmi moderní a aktuální je aplikace výsledků teorie extrémních hodnot v pojišťovnictví, financích a v životním prostředí. Sociální statistika nabývá na významu, protože je hlavním zdrojem informací o státním hospodařství.

Malý příklad: Ve spolupráci s meteorology jsme porovnávali chování maximálních denních teplot v letních měsících v období 1961-1990 ve stanicích na jižní Moravě (např. Kuchařovice, 48°53'N, 16°05'E) s teplotními výstupy globálního cirkulačního modelu ECHAM3/42, generovanými pro stejné zeměpisné souřadnice a pro stejná data. Otázka byla, zda generátor počasí vytváří časovou řadu teplot odpovídající skutečné situaci. Modelem takové řady je tzv. autoregresní model, jehož řád udává, kolik dní si počasí na Moravě „pamatuje“. Bylo třeba sestavit statistický test hypotézy, že generovaná i skutečná časová řada mají stejné autoregresní řady proti alternativě odlišných řádů, bez předpokladu normálního rozdělení pravděpodobností, které zde nemá opodstatnění. Aplikací testu na zmíněná data jsme dospěli k závěru, že řád skutečných maximálních teplot na Moravě je 1, zatímco generátor ECHAM vytváří časovou řadu s řádem 2.



Jiným testem, spolu s Belgičany, jsme ověřovali stupeň závislosti úmrtnosti na kardio-vaskulární choroby v závislosti na maximálních denních teplotách v Bruselu v období 1980-1989, požádání kolegy z Institutu hygieny a epidemiologie v Bruselu. Máme test i výpočetní programy.

Vzhledem k častým povodním v poslední době je třeba vhodně modelovat časovou řadu srážek, a následně zkonstruovat vhodné statistické postupy. O tom diskutujeme s kolegy z aténské univerzity, kteří mají k dispozici podrobná data srážek, jak v čase tak v prostoru – obtížný problém.