



ÚVOD DO MATEMATICKÉ BIOLOGIE I. (setkání čtvrté)



prof. Ing. Jiří Holčík, CSc.

**UKB, pav. A29, RECETOX, dv.č.112
holcik@iba.muni.cz**

© Institut biostatistiky a analýz

MENDLOVY OHLASY

Wilhelm O.Focke (1834 – 1922)

Die Pflanzen-Mischlinge, Ein Beitrag zur Biologie der Gewächse (1881)

„Mendel se domnívá, že našel konstantní číselné poměry mezi typy kříženců.“

1900

Hugo de Vries (Amsterdam)

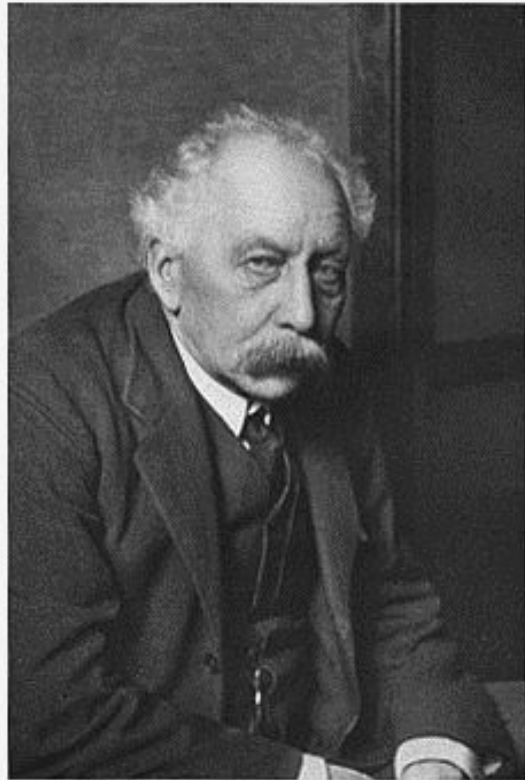
Carl Correns (Tübingen)

Erich von Tschermak (Wien)

1901 - William Bateson – první překlad Mendelovy publikace do angličtiny

Bateson, W.: *Mendel's Principles of Heredity*. Oxford Univ. Press, 1913

WILLIAM BATESON



W. Bateson

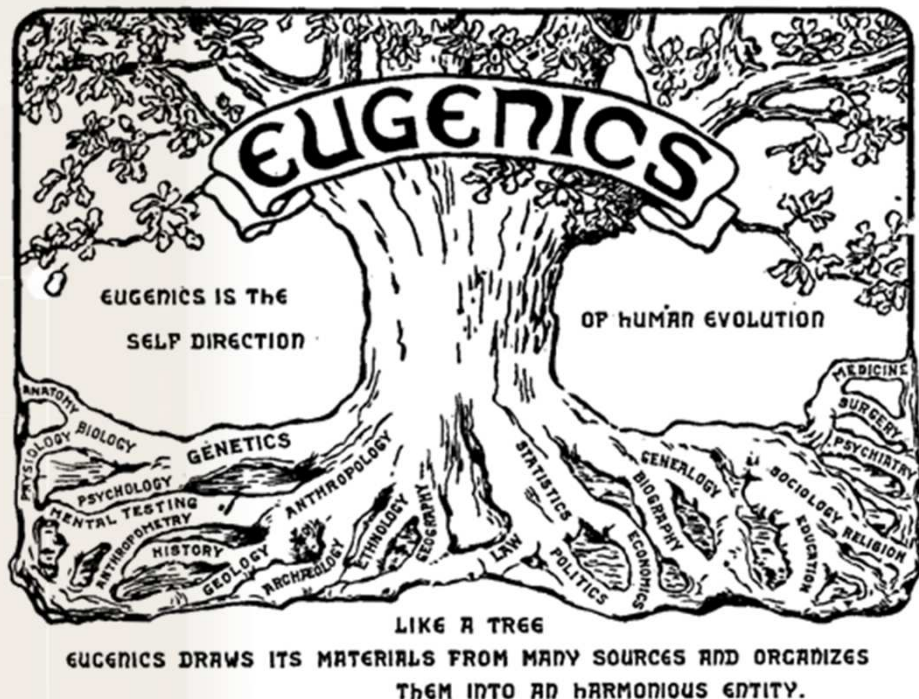
(1861 – 1926)

anglický biolog, hlavní propagátor myšlenek J.G.Mendela po znovu objevení jeho práce 1900

- ✓ první překlad Mendelovy práce do angličtiny (1901)
- ✓ první použil slova **genetika** pro popis vědy o dědičnosti

„Jediný způsob jak dospět k pravdivému výkladu mechanismů dědičnosti je systematická příprava a realizace rozmnožovacích experimentů, tj. výzkum, který asi vyžaduje více trpělivosti a více prostředků než jakékoliv jiné biologické zkoumání.“

EUGENIKA

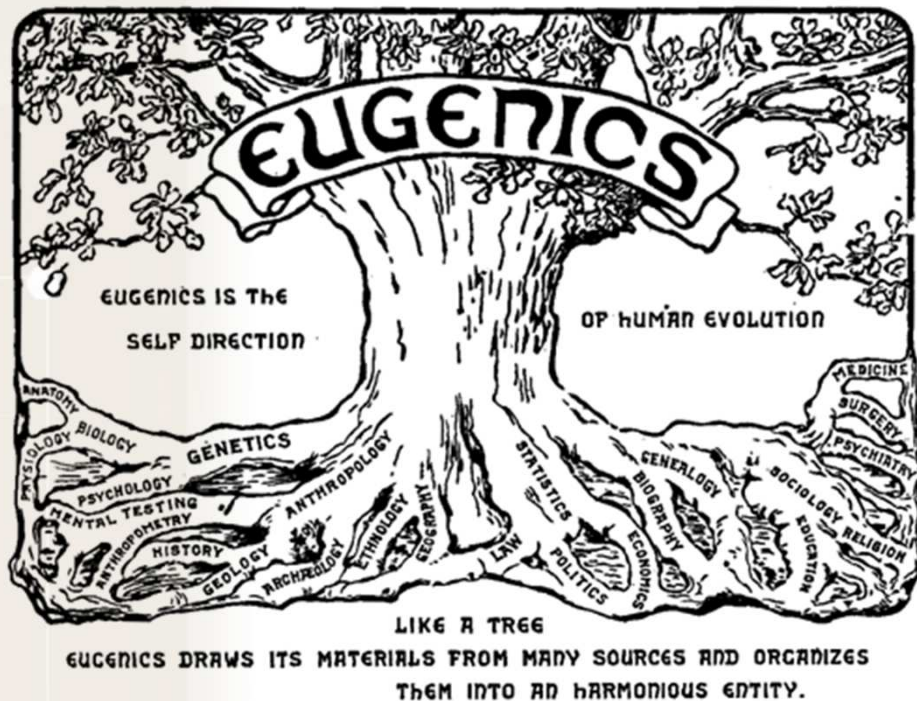


z řečtiny εὐγενής „dobře narozený“
z εὖ - „dobrý, dobře“ a γένος -
„rasa, rod, plemeno, příbuzenstvo“

soubor názorů a praktik,
jejichž cílem je zlepšení
kvality lidské populace

pozitivní a negativní
eugenika

EUGENIKA



z řečtiny εὐγενής „dobře narozený“
z εὖ - „dobrý, dobře“ a γένος -
„rasa, rod, plemeno, příbuzenstvo“

soubor názorů a praktik,
jejichž cílem je zlepšení
kvality lidské populace

pozitivní a negativní
eugenika



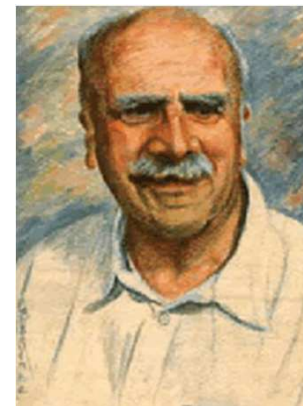
Francis GALTON
1822-1911



Karl PEARSON
1857-1936



W.F. Raphael WELDON
1860-1906



John B. S. HALDANE
1892-1964



Ronald A. FISHER
1890-1962

Sir FRANCIS GALTON, F.R.S.



anglický viktoriánský polyhistor
zeměpisec, meteorolog, cestovatel a
průzkumník tropických oblastí,
zakladatel diferenciální psychologie a
psychometriky, vynálezce identifikace
podle otisků prstů,

bylostný zastánce teorie dědičnosti, protogenetik,
progresivista, eugenik, průkopník statistické korelace
a regrese

- ❖ **regrese k průměru** \Rightarrow souvislosti mezi náhodnými veličinami \Rightarrow **regresní analýza**

Sir FRANCIS GALTON, F.R.S.

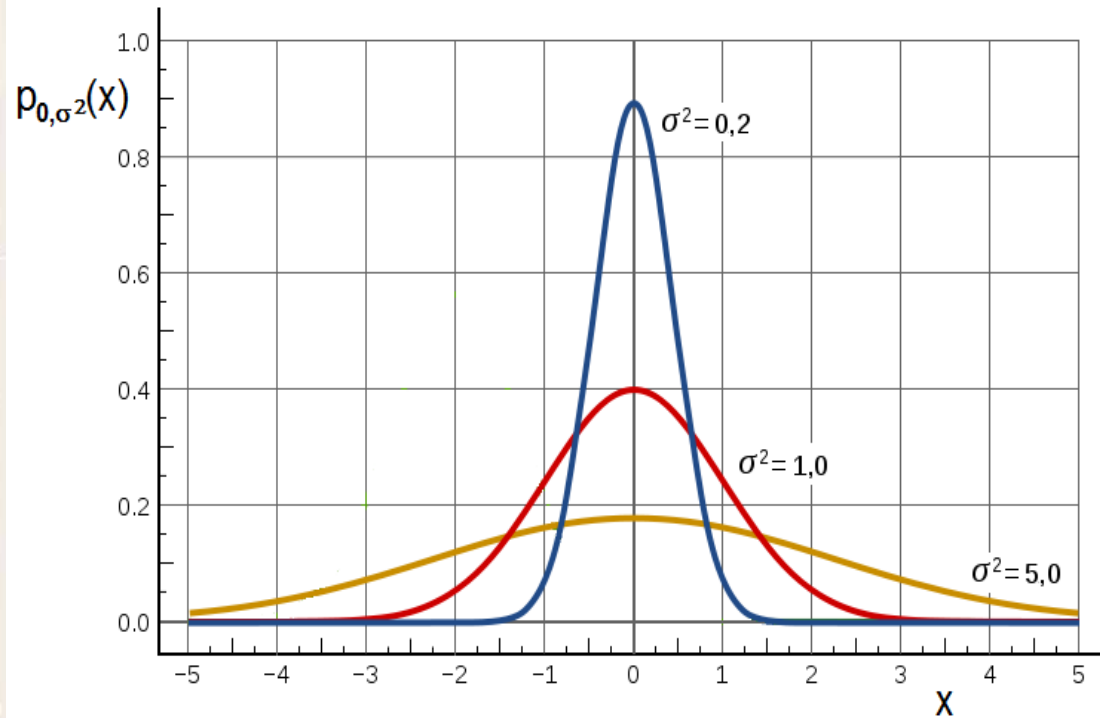


anglický viktoriánský polyhistor
zeměpisec, meteorolog, cestovatel a
průzkumník tropických oblastí,
zakladatel diferenciální psychologie a
psychometrie, vynálezce identifikace
podle otisků prstů,

bylostný zastánce teorie dědičnosti, protogenetik,
progresivista, eugenik, průkopník statistické korelace
a regrese

- ❖ **rozptyl** - popis odchylek kolem střední hodnoty \Rightarrow
experimentální stanovení normálního rozložení

NORMÁLNÍ ROZLOŽENÍ



$$p_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{\pi}} e^{-(x-\mu)^2/\sigma^2}$$

$$\mu = \frac{1}{N} \sum_{n=1}^N x_n \quad \sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - \mu)^2}$$

CENTRÁLNÍ LIMITNÍ VĚTA

Centrální limitní věta v teorii pravděpodobnosti označuje tvrzení, podle něhož se (za určitých podmínek) rozdělení výběrového průměru po vhodné normalizaci blíží k normálnímu rozdělení. O náhodné veličině s uvedeným chováním říkáme, že má *asymptoticky normální rozdělení*.

Ljapunovova věta

Rozdělení součtu vzájemně nezávislých veličin X_i konverguje k normálnímu rozdělení i v případě, že veličiny X_i nemají stejné rozdělení pravděpodobnosti.

BIOSTATISTIKA ☯ BIOMETRIE

aplikace statistiky pro řešení problémů biologie (včetně lékařství, farmacie, zemědělství či rybolovu) -

- ☑ návrh experimentů;
- ☑ sběr, sumarizace a analýza experimentálních dat;
- ☑ interpretace a vyvozování závěrů ze získaných výsledků.



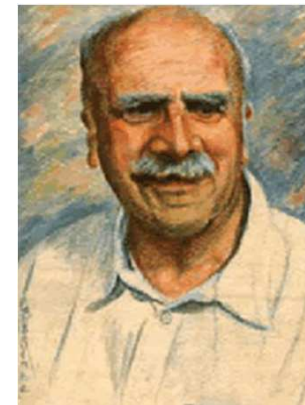
Francis GALTON
1822-1911



Karl PEARSON
1857-1936



W.F. Raphael WELDON
1860-1906



John B. S. HALDANE
1892-1964



Ronald A. FISHER
1890-1962

BIOSTATISTIKA ☯ BIOMETRIE

aplikace statistiky pro řešení problémů biologie (včetně lékařství, farmacie, zemědělství či rybolovu) -

- ☑ návrh experimentů;
- ☑ sběr, sumarizace a analýza experimentálních dat;
- ☑ interpretace a vyvozování závěrů ze získaných výsledků.



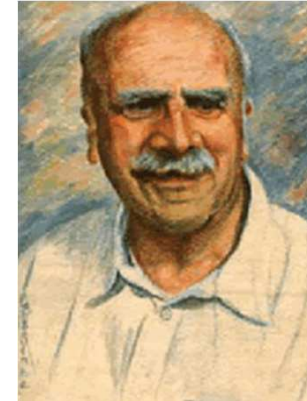
Sewall WRIGHT
1889-1988



Karl PEARSON
1857-1936



W.F.Raphael WELDON
1860-1906



John B. S. HALDANE
1892-1964



Ronald A. FISHER
1890-1962

WALTER FRANK RAPHAEL WELDON DSc FRS



zoolog – morfologie obratlovců -> problémy organických variací -> kontakt s Galtonem a použití jím vyvinutých statistických nástrojů

„problém evoluce živočichů je bytostně statistický problém“

spolupráce s matematikem Carlem Pearsonem -> seznámení s Galtonem

Weldon, Galton, Pearson – projekt Royal Society Committee „For the Purpose of Conducting Statistical Enquiry into the Variability of Organisms“

„otázky vyplývající z Darwinovy hypotézy jsou čistě statistické a statistická metoda je v současnosti zřejmě jediná, jíž lze tyto hypotézy experimentálně ověřit“

KARL PEARSON, F.R.S.

právník, germanista, eugenik

zakladatel matematické statistiky

- roce 1911 založil na University

College London první na světě

katedru statistiky - (vliv na

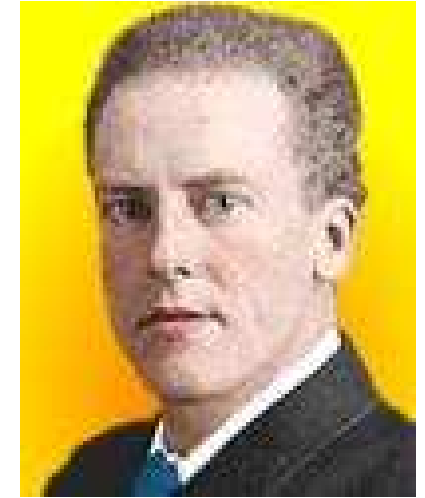
biometriku, meteorologii, teorii

sociálního darwinismu)

chránělec a životopisec Francise Galtona

Pearsonův korelační koeficient

Pearsonův χ^2 test (test dobré shody)



KORELACE



Ve specifičtějším slova smyslu se pojem korelace užívá ve **statistice**, kde znamená vzájemný lineární vztah mezi znaky či veličinami X a Y .

Tento vztah může být kladný, pokud (přibližně) platí $y = kx$, nebo záporný ($y = -kx$).

KORELACE



Ve specifičtějším slova smyslu se pojem korelace užívá ve **statistice**, kde znamená vzájemný **lineární** vztah mezi znaky či veličinami X a Y .

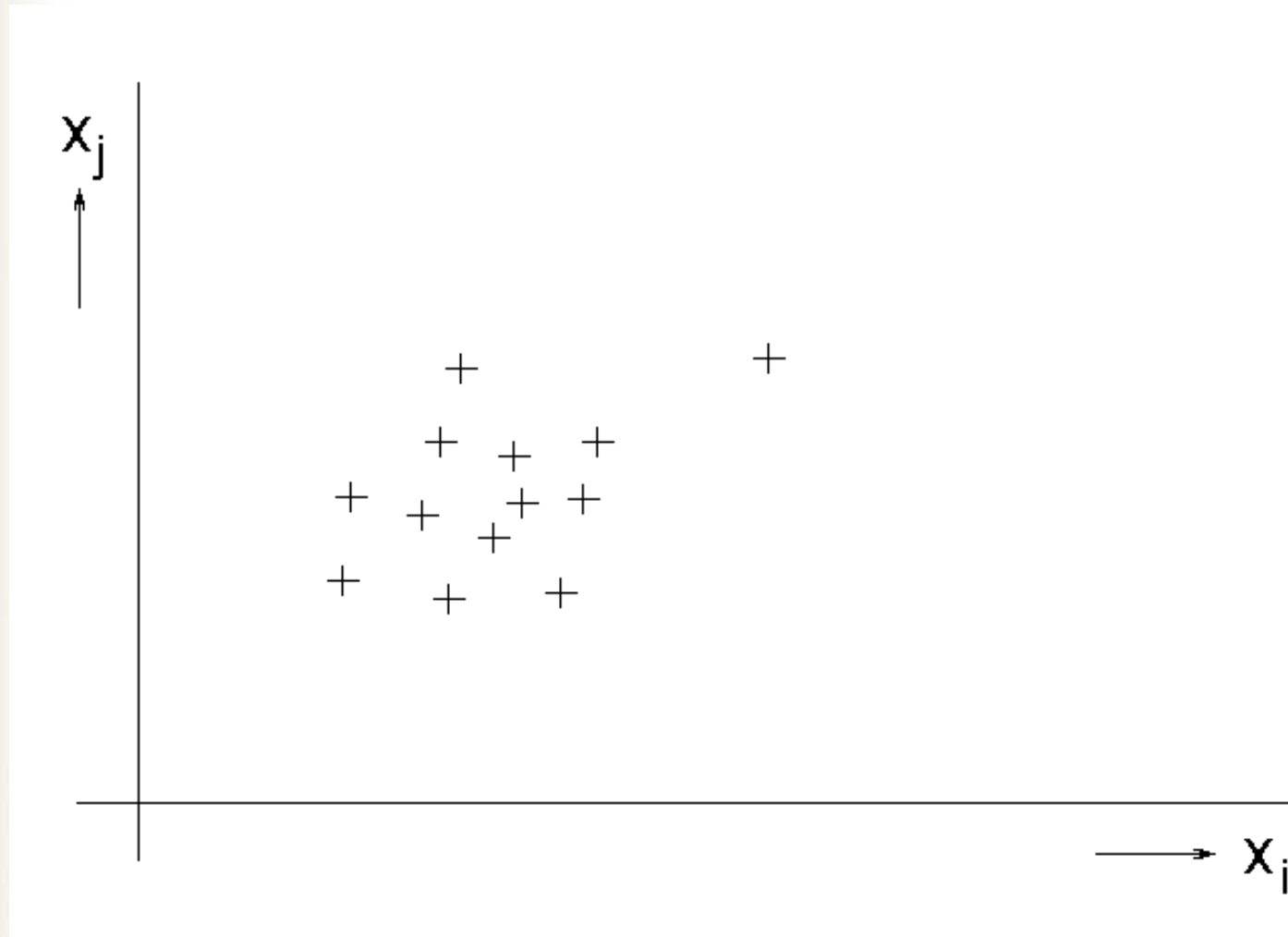
Tento vztah může být kladný, pokud (přibližně) platí $y = kx$, nebo záporný ($y = -kx$).

!!! Pozor !!!

KORELACE NENÍ KAUZALITA

!!! Pohov !!!

KORELACE



KORELACE



Ve specifičtějším slova smyslu se pojem korelace užívá ve **statistice**, kde znamená vzájemný **lineární** vztah mezi znaky či veličinami **x** a **y** .

Tento vztah může být kladný, pokud (přibližně) platí $y = kx$, nebo záporný ($y = -kx$).

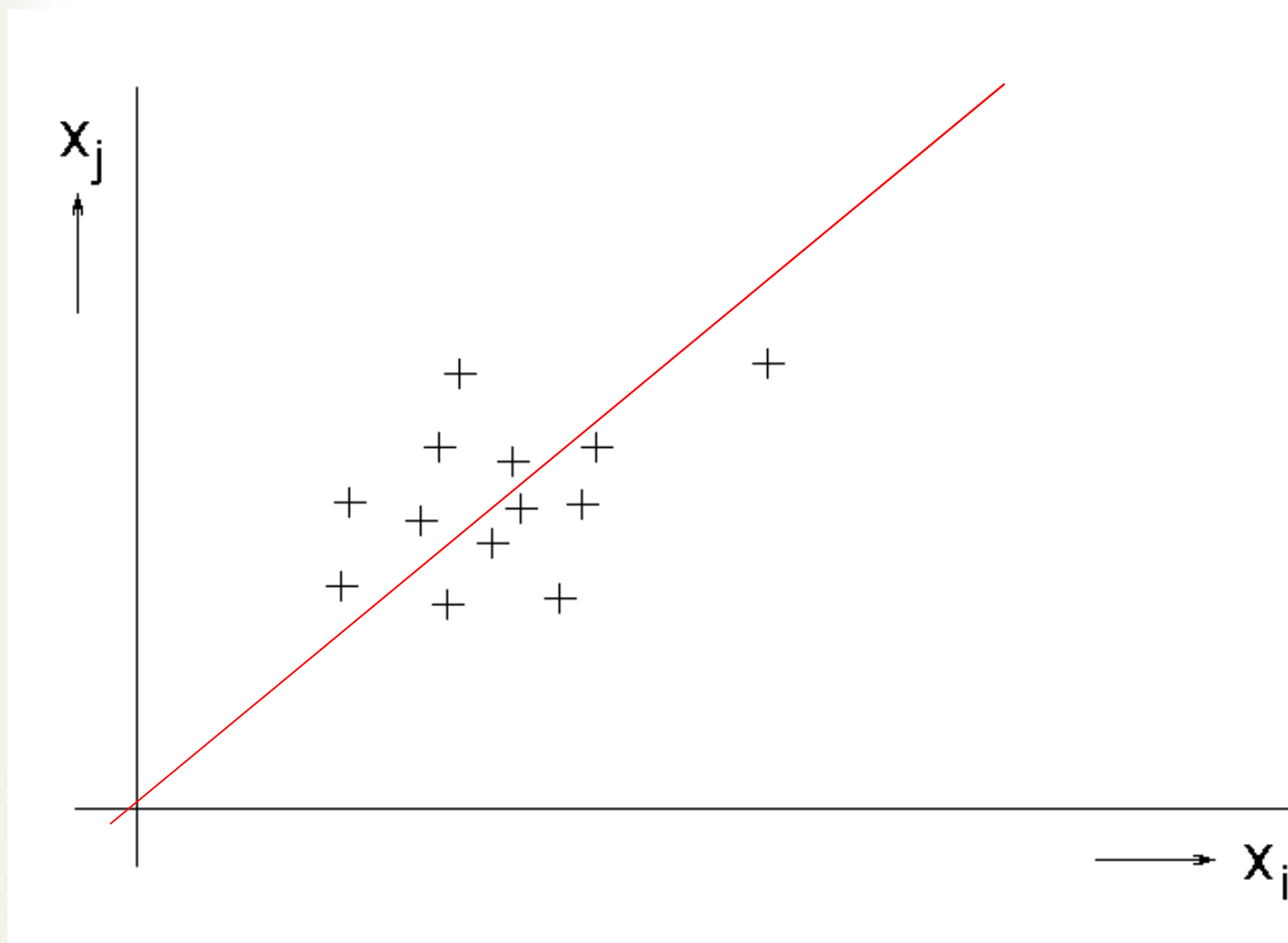
Míru korelace v tom případě vyjadřujeme **korelačním koeficientem**.

PEARSONŮV KORELAČNÍ KOEFICIENT

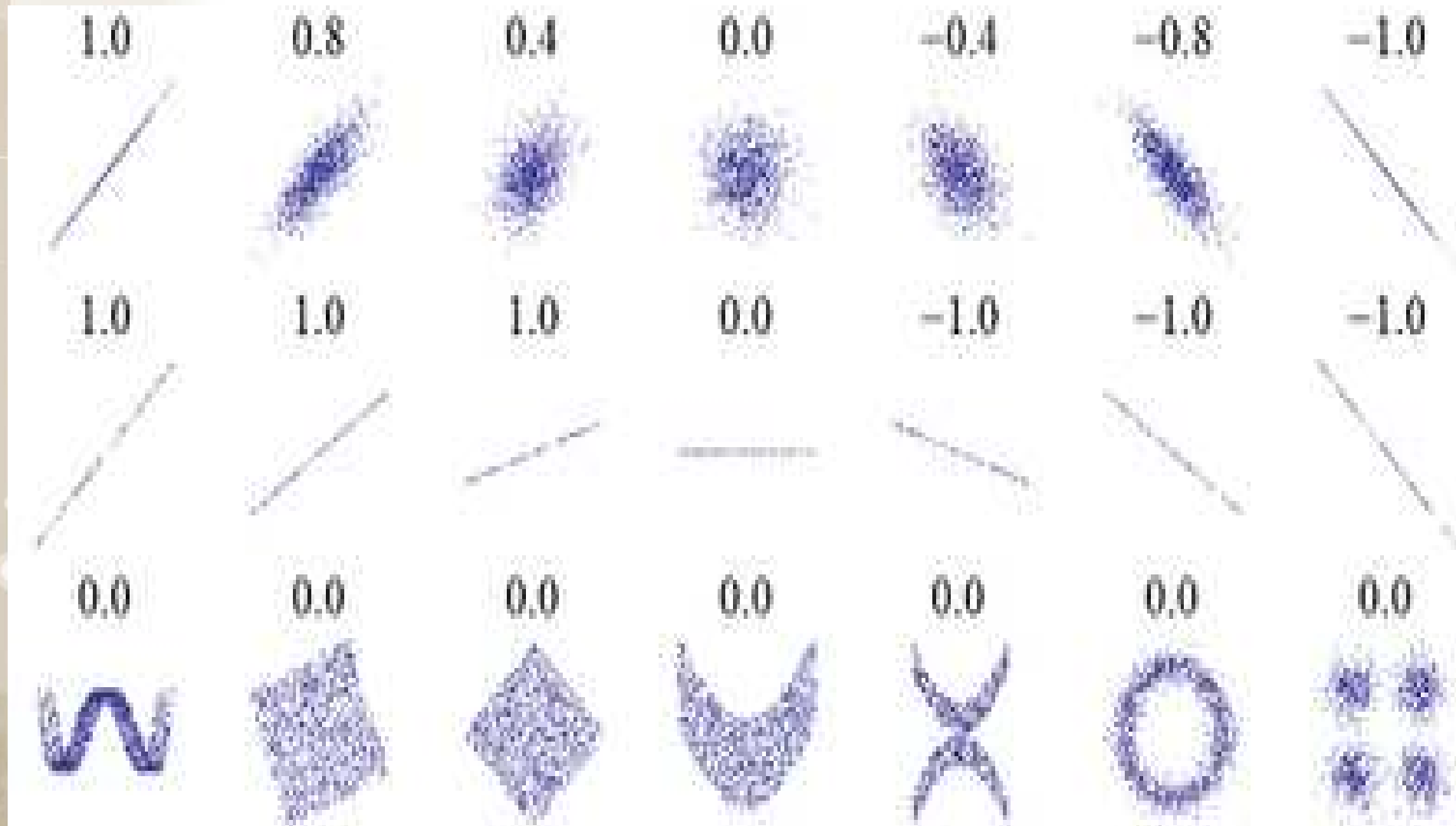
$$\begin{aligned}\rho_{X,Y} &= \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y} = \\ &= \frac{\frac{1}{N} \sum_{n=1}^N (x_n - \mu_X)(y_n - \mu_Y)}{\sigma_X \sigma_Y}\end{aligned}$$

$$\mu_X = \frac{1}{N} \sum_{n=1}^N x_n \quad \sigma_X = \sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - \mu_X)^2}$$

KORELACE



PEARSONŮV KORELAČNÍ KOEFICIENT



PEARSONŮV KORELAČNÍ KOEFICIENT

- ☑ nabývá hodnot od -1 do $+1$, které značí perfektní lineární vztah (záporný nebo kladný)
 - v případě kladné korelace hodnoty obou proměnných zároveň stoupají;
 - v případě záporné korelace hodnota jedné proměnné stoupá a druhé klesá;
 - v případě neexistence lineárního vztahu $r = 0$;
- ☑ je nezávislý na jednotkách původních proměnných, je bezrozměrný;
- ☑ při změně pořadí proměnných se výše korelačního koeficientu nemění;
- ☑ korelační koeficient je platný pouze v rozmezí daném použitými daty;
- ☑ korelační koeficient výrazně odlišný od nuly není důkazem funkčního vztahu proměnných, jiného než lineárního;
- ☑ malá hodnota korelačního koeficientu není známkou nefunkčního vztahu proměnných.

WALTER FRANK RAPHAEL WELDON DSc FRS



WELDONOVY KOSTKY

1894 experiment s 12 hracími kostkami 26306 vrhů
– chtěl ověřit, zda jsou či nejsou rozdíly mezi četnostmi v jednotlivých sériích a teoretickým pravidlem uvažovaným v celku, větší než lze přisoudit fluktuacím náhodného vzorkování



data použita Pearsonem pro základní práci o χ^2 statistice

Pearson, K.: On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that It Can Be Reasonably Supposed to Have Arisen from Random Sampling. *Phil. Mag.*, 1(1900):157-175.

Zacariah Labby (2009). "Weldon's Dice, Automated". *Chance*. **22** (4): 6–13.

PEARSONŮV χ^2 TEST

(TEST DOBRÉ SHODY)

- ❖ používáme obecně k testování shody četnosti (především u nominálních znaků – kategoriálních dat), ale i znaků kvantitativních.
- ❖ je založen na posouzení rozdílu mezi skutečnými (empirickými) četnostmi výskytu hodnot ve výběrovém souboru a očekávanými (teoretickými) četnostmi odpovídajícími příslušnému předpokládanému rozdělení pravděpodobností.
- ❖ test rozhoduje, zda je rozdíl mezi empirickými a teoretickými četnostmi způsoben pouze náhodně, nebo je rozdíl natolik veliký, že je způsoben tím, že výběrový soubor pochází z populace odpovídající nějakému jinému neznámému rozdělení.

WALTER FRANK RAPHAEL WELDON DSc FRS



Weldon, W.F.R.: Mendel's Laws of Alternative Inheritance in Peas.
Biometrika. 1:228-254.

zákon dominance:

Pokud jsou zkříženy dvě rostliny hrachu různých druhů, pak hybridní potomek bude vykazovat pouze dominantní vlastnost rodičů. Bude ji vykazovat (nebo téměř) beze změn. Recesivní vlastnost se neprojeví, resp. tak málo, že uniká pozornosti.

zákon segregace:

Pokud jsou kříženci první generace dále kříženy, pak se v druhé generaci projeví všechny zděděné kombinace vlastností, přičemž se budou řídit zákonem dominance. Smíšené vlastnosti se neobjeví.

ROSTLINY S DOMINANTNÍM ZNAKEM V 2. GENERACI HYBRIDŮ

	rostlin v 2. generaci	počet dominantních znaků	teoretický počet dominantních znaků	teoretická chyba ε	odchylka pozorování od teoretického počtu
tvár semen	7324	5474	5493	$\pm 24,995$	-19
zabarvení dělohy	8023	6022	6017,25	$\pm 26,160$	+4,75
barva květu (semene)	929	705	696,75	$\pm 8,902$	+8,25
tvár lusku	1181	882	885,75	$\pm 10,037$	-3,75
barva lusku	580	428	435	$\pm 7,034$	-7
umístění květů	858	651	643,5	$\pm 8,555$	+7,5
velikost stonku	1064	787	798	$\pm 9,527$	-11

ČÁST HYBRIDŮ V 2. GENERACI S DOMINANTNÍM ZNAKEM, KTERÉ PŘENESLY DOMINANTNÍ ZNAK NA POTOMKY

	počet sledovaných hybridů v 2. generaci	počet přenesených pouze dominantních znaků	teoretický počet	teoretická chyba ε	odchylka pozorování od teoretického počtu
tvar semen	565	193	188,3	$\pm 7,558$	+4,7
zabarvení dělohy	519	166	173	$\pm 7,244$	-7,0
barva květu (semene)	100	36	33,3	$\pm 3,180$	+2,7
tvar lusku	100	29	33,3	$\pm 3,180$	-4,3
barva lusku	100	40	33,3	$\pm 3,180$	+6,7
umístění květů	100	33	33,3	$\pm 3,180$	-0,7
velikost stonku	100	28	33,3	$\pm 3,180$	-5,3

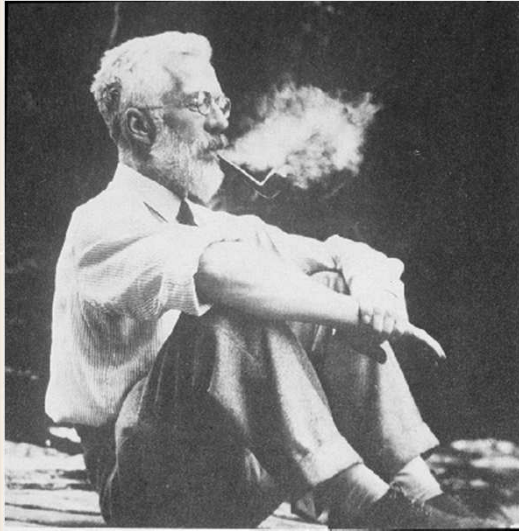
WALTER FRANK RAPHAEL WELDON DSc FRS



Weldon, W.F.R.: Mendel's Laws of Alternative Inheritance in Peas.
Biometrika. 1:228-254.

- ❖ fialové skvrny, které se objevují na povrchu semene jsou u kříženců větší a častější než u „čistokrevných“ forem;
- ❖ jako nevýznamné se jeví (dříve jinými pozorované), že křížení způsobuje zvětšení vegetativních orgánů, takže kříženci jsou často vyšší než jejich rodiče.

Sir RONALD AYLMER FISHER, F.R.S.



anglický statistik, evoluční biolog,
matematik, genetik a eugenik

„genius, který téměř sám vytvořil
základy moderní statistiky“

Richard Dawkins: „největší biolog od
dob Darwina“

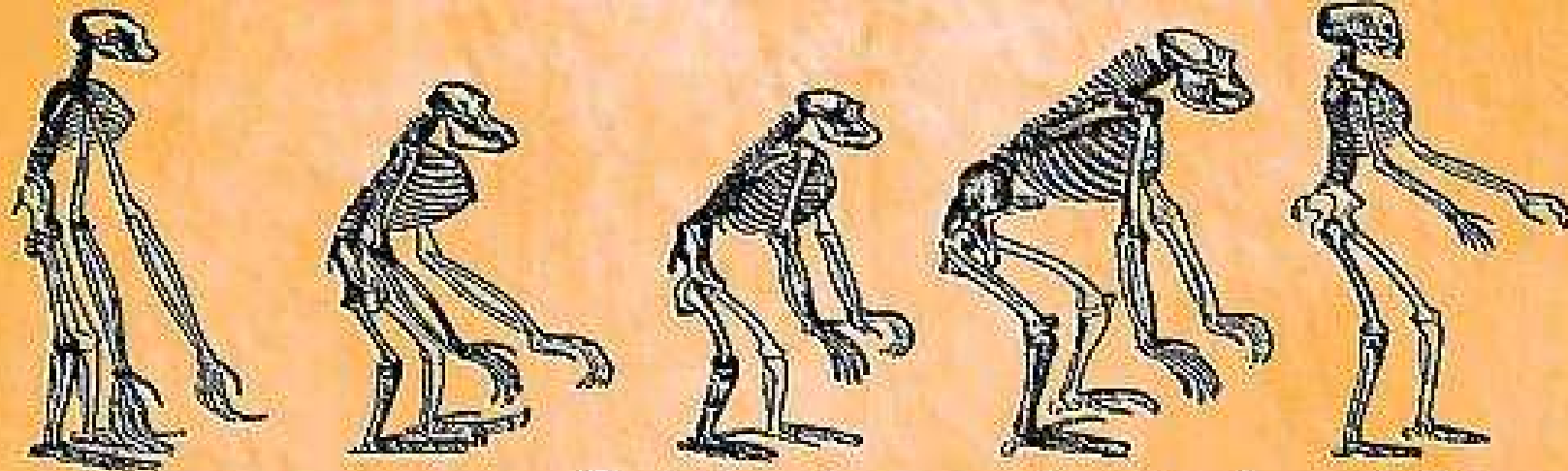
Geoffrey Miller:

„**pro biology** – architekt moderní syntézy, který použil
matematické modely pro integraci Mendlový genetiky a
Darwinovy teorie vývoje“

„**pro psychology** – vynálezce různých dosud používaných
statistických testů“

„**pro zemědělce** – zakladatel experimentálního výzkumu, který
zachránil milióny lidí před hladem zavedením racionálních
programů křížení obilí“

Sir RONALD AYLMER FISHER, F.R.S.



Theorem of Natural Selection
The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.

Ronald A. Fisher



More science quotes at Today in Science History — todayinsci.com

Sir RONALD AYLMER FISHER, F.R.S.

Natural selection is a mechanism for generating an exceedingly high degree of improbability.

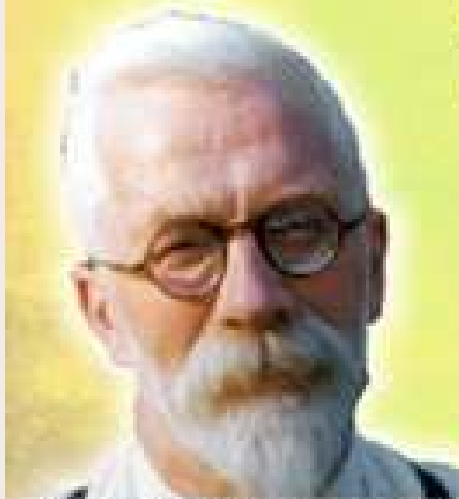
Sir Ronald Aylmer Fisher



More science quotes at Today in Science History todayinsci.com

Sir RONALD AYLMER FISHER, F.R.S.

No isolated experiment, however significant in itself, can suffice for the experimental demonstration of any natural phenomenon.



Sir Ronald Aylmer Fisher

More science quotes at Today in Science History todayinsci.com

Sir RONALD AYLMER FISHER, F.R.S.

matematické výdobytky:

- ✓ 1919 analýza rozptylu – **ANOVA**
- ✓ 1912 – 1922 vývoj, analýza a popularizace principu maximální věrohodnosti
- ✓ sjednocení principů testů statistických hypotéz – Studentův t- test + χ^2 test



1925 „Statistical Methods for Research Workers“

1930 „The Genetical Theory of Natural Selection“

1935 „The Design of Experiments“

1938 spolu s Frankem Yatesem „Statistical Tables for Biological, Agricultural, and Medical Research“

1936 – „Has Mendel’s Work Been Rediscovered?“
Annals of Science, 1(2):115-126.

Sir RONALD AYLMER FISHER, F.R.S.

1936 – „Has Mendel's Work been Rediscovered?“ *Annals of Science*, 1(2):115-126.



Co Mendel objevil?

Jak to objevil?

Co si myslel, že objevil?

Jaký význam měly jeho objevy pro vědu jeho doby a jaká byla reakce na ně?

Měl by Mendel být brán doslova?

„Test významnosti odchylek od očekávaných hodnot v binomických řadách byly matematikům známy při nejmenším od poloviny 18. století. Na druhé straně jeho studium matematiky ve Vídni se problematikou teorie pravděpodobnosti zabývalo nepříliš.“

Sir RONALD AYLMER FISHER, F.R.S.

1936 – „Has Mendel's Work been Rediscovered?“ *Annals of Science*, 1(2):115-126.



Mendelovy výsledky jsou konzistentně tak blízké očekávaným, že jsou podezřelé

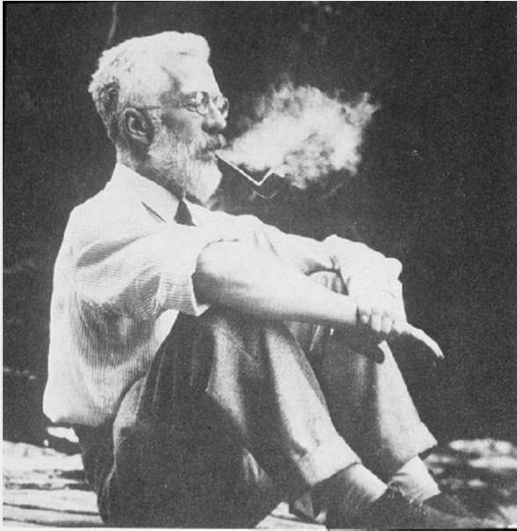
Mendel pravděpodobně chybně klasifikoval některé heterozygoty jako homozygoty při stanovení genotypického poměru 2:1.

S. Wright: Mendel nevědomě klasifikoval některé jedince s neurčitým fenotypem ve prospěch jeho očekávání

POCHYBNOSTI O MENDELOVĚ PRÁCI

- ☑ nejsou Mendelova data příliš dobrá, aby byla pravdivá?
- ☑ není popis jeho experimentů pouze fiktivní?
- ☑ vyslovil zákony dědičnosti, které se mu připisují?
- ☑ rozpoznal Mendel, ale nezmínil dědičné (chromozomové) vazby?
- ☑ podporoval nebo oponoval Darwinovi?

Sir RONALD AYLMER FISHER, F.R.S.



Doll, Richard, Hill A.Bradford: Smoking and Carcinoma of the Lung. Preliminary Report, Br.Med.J. (1950), p.739 - 748

Levin Morton L.,Goldstein Hyman, Gerhardt Paul R.: Cancer and Tobacco Smoking. A Preliminary Report, *JAMA* (1950), p.336-338.

Fisher R.A.: Alleged Danger of Cigarette-Smoking. Br.Med.J. (1957), p.297 - 298

Fisher R.A.: Cigarettes, Cancer, and Statistics. Centennial Review. (1958), p.151-166.

Fisher R.A.: Lung Cancer and Cigarettes. *Nature* (1958), 108-109.

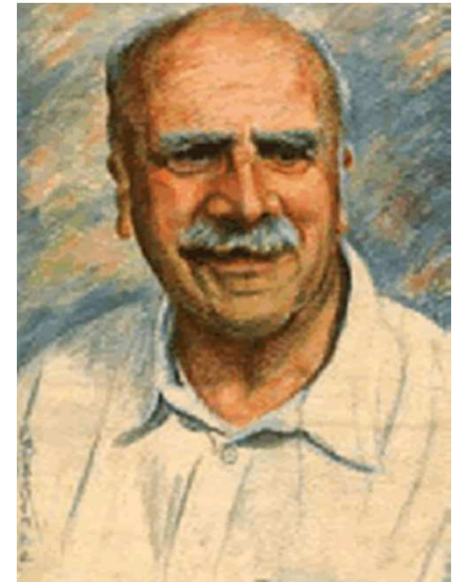
Fisher R.A.: Cancer and Smoking. *Nature* (1958), 596-597.

J.B.S. HALDANE

- ✓ britský/indický biolog/biostatistik
socialista/marxista/ateista

Arthur C. Clark: „zřejmě nejskvělejší popularizátor vědy ve své generaci“

Peter Medawar: „nejchytřejší člověk, kterého jsem kdy potkal“



- ✓ jeden ze zakladatelů matematické teorie (Fisher, Wright) populační genetiky, sjednocující Mendelovu genetiku a Darwinovu evoluci - **neodarwinismus**

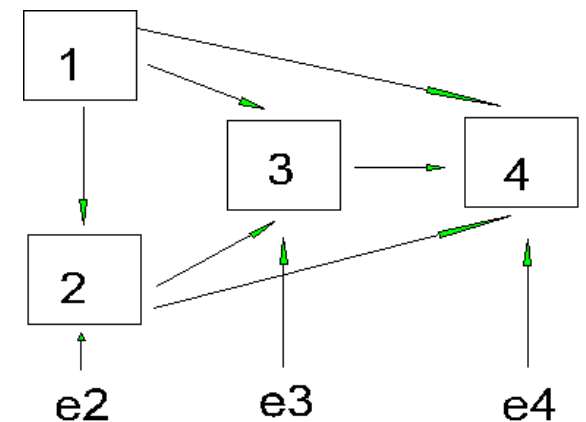
1929 – v článku „Abiogenesis“ zavedl teorii „pravěké polévky“ – základ chemického původu života

- ✓ první přišel s myšlenkou umělého oplodnění, klonování

SEWALL WRIGHT



- ✓ americký genetik
- ✓ spoluzakladatel populační genetiky (evoluční teorie)
- ✓ základní práce v oblasti „path analysis“ (1921) – výpočet a popis orientovaných závislostí mezi množinou proměnných
- ✓ definoval koeficient křížení a způsob jeho výpočtu - míra věrohodnosti genetických efektů způsobených křížením a očekávaných na základě znalosti rodokmenu



NA SHLEDANOU ZA 2 TÝDNY