O ZÁKONU POČÁTEČNÍ HODNOTY

Jan Klaschka, Výzkumný ústav psychiatrický, Praha (1988),
Psychiatrické centrum, Praha (1991)

Tato práce budiž chápána jako spíše osvětová než teoretická. Se situací, již zde budeme studovat, se pravděpodobně setká dříve či později každý, kdo přichází do styku, byť i na elementární úrovni, s analýzou biomedicínských dat. Autorovi se přitom zdá, že není příliš vyvinuté obecné povědomí o (celkem triviálním) nebezpečí, které s sebou takové setkání nese, a že tedy má dobrý smysl upozornit na danou situaci jako na úsek častých vědeckých nehod.

Přistupme k našemu tématu - nejlépe oklikou. V 5. čísle PMFA v r. 1984 jsme si ve fejetonu Iana Stewarta [7] mohli přečíst tuto pasáž:
"... Právě jsem si vzpomněl na Případ drogovaných paviánů, kteří reagovali na daný podnět stlačením páčky v čase D ve srovnání s dobou reakce C u nedrogovaných paviánů. Desetiletý výzkum vedl pouze k vyznačení hodnoty D/C vedle C na logaritmické stupnici, přičemž se pomocí počítače obdržela korelace blízká hodnotě -1; to bylo interpretováno jako závislost účinků drogy na kontrolní době reakce C. Potom si v roce 1978 někdo povšiml, že když

$$\log (D/C) \sim -\log C + k \qquad (k = \text{konstanta}), \tag{1}$$

pak

$$D \sim e^k = \text{konstanta}.$$

Kdybych já byl udělal paviánům totéž, co udělali doktoři mé milované matematice, byl bych teď ve vězení. (Copak si nikdo předtím nepovšiml, že D bylo zhruba konstantní?...)"

Tato zpráva z černé kroniky aplikací matematiky je sice zábavná, ale zcela jasná je zřejmě jen non-statistikům. Pro nás ostatní není bez otazníků: Jaké bylo uspořádání experimentu? Co přesněji znamená "korelace blízká hodnotě -1"? Jakým způsobem se z korelace blízké -1 stal regresní koeficient blízký -1?

Pokusme se (bez nároku na neomylnost) zprávu vyložit po svém: U všech pokusných zvířat byla jednou změřena doba reakce C před podáním drogy a doba D po podání. Jak tomu v podobných experimentech bývá, rozdělení hodnot veličiny C resp. D bylo šikmé, podobné spíše logaritmicko-normálnímu než normálnímu. Proto byly statistické analýze klasickými metodami místo samotných veličin C a D podrobeny veličiny $X_1 = \log C$ a $X_2 = \log D$. Párovým t-testem X_2 proti X_1 bylo zřejmě ověřeno, že droga způsobuje systematické zpomalení (nebo zrychlení) reakce. Následovala logická otázka, zda účinek drogy souvisí s vlastnostmi zvířete známými před jejím podáním, speciálně s dobou reakce v normálním stavu. Nic nebylo přirozenější, než vyšetřit korelační, resp. lineární regresní vztah mezi rozdílem $X_2 - X_1$ (= log (D/C)) chápaným jako míra účinku drogy a veličinou X_1 (= log C), která reprezentuje normální stav. (Tak se ve vztahu (1) objevila obzvláštním šarlatánstvím zavánějící kouzla s logaritmy.) Zřejmě byl vypočten záporný a od nuly signifikantně odlišný kerelační koeficient (popř. lineární koeficient regresní rovnice, což co do znaménka a signifikance vyjde nastejno) - výklad pak už známe z citovaného úryvku.

Mohlo by se zdát, že se vlivem našeho nového výkladu Případ drogovaných paviánů začíná nenápadně ocitat mimo černou kroniku. Z toho ovšem bude čtenář za chvíli vyveden.

Výše uvedený výsledek týkající se paviánů, drogování a (logaritmu) doby reakce je jen jedním prvkem obsáhlé třídy W výsledků biomedicínského výzkumu, jejíž prvky můžeme symbolicky označit W(A, B, X). A zde značí nějaký typ živých objektů, B nějaký zásah a X nějakou veličinu měřenou u daných objektů.

Výsledky W(A, B, X) jsou charakterizovány jednak fakty, jednak jejich interpretací. Výchozím faktem je zde platnost vztahu

$$\mathrm{cor}\ (X_2 - X_1, X_1) < 0 \qquad (2)$$

(ověřená, pochopitelně s určitou dávkou nejistoty, statistickým testem) pro dvojici měření X_1, X_2 veličiny X před zásahem a po něm. Charakteristická interpretace vztahu (2) pak je, že u objektů typu A účinek zásahu B na míru vlastnosti korespondující s veličinou X závisí na míře dané vlastnosti před zásahem, přičemž závislost je typu "čím větší počáteční hodnota, tím menší efekt spočívající ve zvýšení, resp. tím větší efekt spočívající ve snížení".

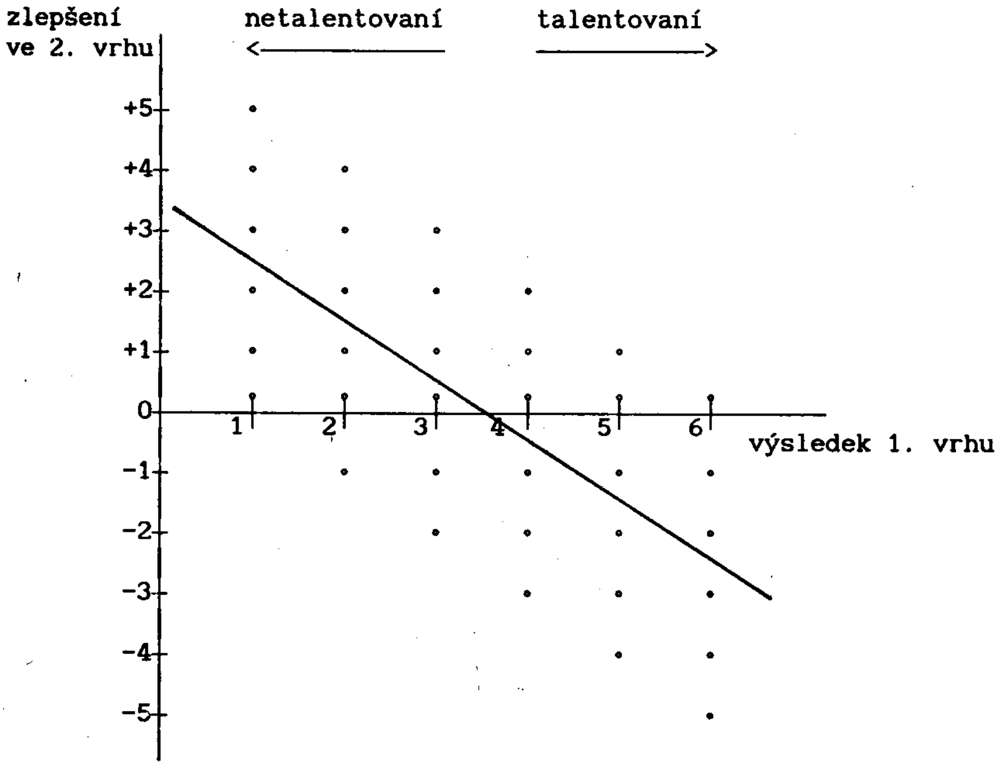
Ve třídě W bychom našli práce dokazující např., že určitý lék snižuje krevní tlak tím účinněji, čím je pacient těžší hypertonik, že redukční dieta je účinnější u obéznějších, ale také, že vedlejší účinek léku (měřený zvýšením koncentrace nežádoucí látky v krvi) je méně výrazný u těch pacientů, pro které by byl nejnebezpečnější (tj. u těch, kteří měli vysokou hladinu nežádoucí látky v krvi na začátku léčení).

Označení W pro uvedenou třídu výsledků jsme použili proto, že výsledky W(A, B, X) bývají často svými autory vydávány za důkaz platnosti tzv. Wilderova zákona počáteční hodnoty v dané oblasti, tj. pro danou trojici (A, B, X).

Kde se uvedený zákon vzal: Počátkem 30. let publikoval Wilder článek [9] (následovaný řadou dalších časopiseckých prací a posléze monografií [10]), v němž formuloval (dle jeho názoru) dosud nepovšimnutý biologický zákon - zákon počáteční hodnoty (Ausgangswert-Gesetz). Zákon počáteční hodnoty - spokojíme-li se s volným přetlumočením - vypovídá o vegetativním nervstvu resp. vegetativních orgánech: čím vyšší je stupeň vybuzení resp. činnosti před podnětem, tím menší je schopnost dosáhnout vyššího vybuzení resp. intenzivnější činnosti vlivem dráždivého podnětu, a naopak tím větší je schopnost vlivem tlumivého podnětu stupeň vybuzení resp. činnosti snížit.

Cílem této práce není vyvracet Wilderův zákon ani dokazovat, že Wilder zákon počáteční hodnoty formuloval veden "statistickou bludičkou". Předmětem našeho zájmu jsou výsledky z W a logika jejich tvorby; dovolává-li se nějaká práce Wilderova zákona, může to pro nás mít hodnotu upozornění na zvýšenou pravděpodobnost výskytu nepravostí v práci; zákon počáteční hodnoty nás bude zajímat jen jako (nepovinná) nálepka na práce z W, ne sám o sobě.

Studujme nyní tento myšlený příklad: Mějme (pokud možno větší) soubor pokusných osob. Každého probanda necháme vrhnout kostkou, poté mu dáme vypít kávu, načež ho necháme hodit kostkou podruhé. Výsledky obou vrhů můžeme zachytit grafem, kde na jedné ose vyneseme výsledky 1. vrhu a na druhé zlepšení ve 2. vrhu (tj. rozdíl výsledků 2. a 1. vrhu). Získáme tak zřejmě graf podobný obr.1. Není pochyb, že poznatek W(lidé, dávka kofeinu, výsledky hodu kostkou) by měl plné statistické oprávnění být zařazen do třídy W. Zbývá jen nález interpretovat tak, že účinky kofeinu na schopnost házet dobře kostkou závisí na talentu hráče, měřeném pochopitelně výsledkem vrhu před podáním kávy; netalentovaní



Obr. 1: Závislost výkonnostního růstu v hodu kostkou po dávce kofeinu na talentu

vrhači jsou povzbuzováni k lepším výkonům, výkonnost nadaných naopak prudce klesá (a proto by káva měla být přísně zakázána ve středisku vrcholových vrhcábů).

Nevážnost našeho příkladu má ovšem nepěkné důsledky pro zbytek třídy W: Bez ohledu na odlišnost věcné podstaty i míry serióznosti řešených problémů je každý výsledek W(A, B, X) (včetně toho o kostkách a kofeinu) podepřen týmiž argumenty; jednotlivé výsledky z W říkají něco podstatného (netriviálního) o léčení hypertenze nebo obezity, vedlejších účincích léků, popř. o reakcích drogovaných paviánů jen tehdy, říká-li také náš příklad něco nesamozřejmého o hodu kostkou a kofeinu.

Až potud jsme se zabývali argumenty použitelnými v případné diskuzi se statistickým laikem. S minimem matematického formalismu nahlédneme "bídu třídy W" daleko rychleji.

V celém zbytku textu zachováme značení X_1 a X_2 měření před zásahem a po něm a budeme předpokládat, že náhodný vektor $(X_1, X_2)'$ má konečné druhé momenty a regulární varianční matici

$$\mathrm{var}\,(X_1, X_2)' = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

Ze vztahu

$$\mathrm{cov}\,(X_2 - X_1, X_1) = \sigma_1\sigma_2\,(\rho - \sigma_1 / \sigma_2) \qquad (3)$$

ihned vyplývá, jak cenný je poznatek, že v rámci konkrétního biomedicínského problému platí vztah (2). Pro $\sigma_1 \geq \sigma_2$ je pravá strana rovnice (3) záporná, a tedy vztah (2) splněn, automaticky. Mj. je zřejmé, že bychom zápornou korelaci mezi změnou $X_2 - X_1$ a počáteční hodnotou X_1 dostali i v případě, že X_1 a X_2 mají stejné rozdělení (a tedy $\sigma_1 = \sigma_2$), tj. např. při sledování spontánní intraindividuální variability nějaké veličiny u pokusných objektů, jež nebyly v době mezi měřeními podrobeny vůbec žádnému zásahu.

Úhrnem můžeme konstatovat, že poznatek (2) o veličinách X_1 a X_2 měřených na biologických objektech je jen málo víc poznatkem o živé přírodě, než je jím objev, že dva pelikáni plus dva pelikáni jsou čtyři pelikáni.

- * -

Udělejme si teď malou inventuru: Otázka, zda efekt zásahu závisí na výchozím stavu živého objektu, je jistě legitimní. Zatím jsme dokázali diskvalifikovat vztah (2) jako kritérium rozhodnutí takové otázky - s žádným lepším kritériem jsme však nepřišli.

Přítomnost X_1 v obou výrazech - pro počáteční hodnotu i efekt - vnáší do korelačního vztahu mezi počáteční hodnotu a efektem artefakt, který většinou rozhoduje o záporném znaménku korelačního koeficientu.

Máme tedy problém: Jak testovat, zda efekt zásahu "podstatně" závisí na počáteční hodnotě, nebo jinak, zda v korelačním vztahu mezi efektem a počáteční hodnotou je "ještě něco navíc" mimo už zmíněný artefakt (přitom se pochopitelně neptáme jen na to, jsou-li, či nejsou veličiny X_1 a X_2 nekorelované).

Problém JAK testovat je ovšem nastolen předčasně - ještě nevíme, CO testovat, tj. jak formulovat hypotézu o "podstatné" souvislosti efektu a počáteční hodnoty.

Že s výsledky ze třídy W není něco v pořádku, se ví už dávno. Problematice vztahu počáteční hodnoty a efektu zásahu (kritice výsledků založených na vztahu (2) i hledání jiných kritérií) se věnovala přinejmenším od 50. let řada německých autorů - jmenujme namátkou alespoň články [2], [4], [6], [8]. Všechny starší práce známé pisateli těchto řádek však řešení našehc problému v nejlepším případě naznačují.

Problém CO a JAK se zdá být natolik frekventovaný v aplikacích (a přitom ne tak hluboký), že by bylo s podivem, kdyby nebyl už desítky let uspokojivě vyřešen. Přesto - překvapivě - se autorovi tohoto sdělení dosud podařilo najít jedinou práci podávající řešení, a to článek Berryho a spoluautorů [1] z r. 1984. Řešení je v článku prezentováno jako původní, tj. bez jakéhokoli odkazu na dřívější pokusy (a mimochodem také bez odkazu na Wilderův zákon počáteční hodnoty). Prvenství Berryho se nechce věřit, ale důkladnější rešerše, jež by je vyvrátila, zatím chybí.

Berry [1] předpokládá, že vektor $(X_1, X_2)'$ měření před zásahem a po něm má dvourozměrné normální rozdělení, $E\, X_1 = \mu$, $E\, X_2 = \mu + \Delta$ (dřívější značení a předpoklady stále platí). Rozlišuje se aditivní efekt a tzv. diferenciální efekt zásahu. Přítomnost aditivního efektu je dána nenulovou hodnotou parametru Δ (a testuje se, jak známo, např. párovým t-testem). Je-li přítomen i diferenciální efekt, změna $X_2 - X_1$ "podstatně" závisí na X_1. V souvislosti se vztahem (2) se mluví o zdánlivém diferenciálním efektu.

Formulace nulové hypotézy týkající se diferenciálního efektu je jednoduchá jako Kolumbovo vejce:

$$H_0 : \sigma_1 = \sigma_2.$$

Jak takovou formulaci zdůvodnit? Třeba tímto prostým způsobem: Neexistuje-li žádný efekt (aditivní ani diferenciální), mají X_1 a X_2 stejné rozdělení. Existuje-li pouze aditivní efekt, liší se rozdělení X_1 a X_2 jen posunutím o Δ (tj. $\sigma_1 = \sigma_2$). Existuje-li i diferenciální efekt, liší se rozdělení X_1 a X_2 ještě něčím jiným než posunutím - pak ale vzhledem k normálnímu rozdělení jiné kritérium existence diferenciálního efektu než shoda či neshoda margiálních rozptylů prostě nezbývá. (Berry neargumentuje přesně takto, ale jeho výklad lze také takto pochopit.)

Protože za platnosti H_0 platí

$$E\,(X_2 - X_1 \mid X_1) = \Delta - (1 - \rho)\,(X_1 - \mu),$$

zavádějí autoři [1] upravený efekt

$$Y = X_2 - X_1 + (1 - \rho)(X_1 - \mu).$$

Existence či absence diferenciálního efektu se odrazí ve vztahu Y k X_1:

$$E\,(Y \mid X_1) = \Delta + \rho\,(\sigma_2 / \sigma_1 - 1)\,(X_1 - \mu).$$

Platí-li H_0, je tedy $E\,(Y \mid X_1) = \Delta$, v opačném případě - s výjimkou $\rho = 0$ - je $E\,(Y \mid X_1)$ rostoucí či klesající funkcí argumentu X_1 - smysl trendu závisí na znaménku členu $\rho\,(\sigma_2 / \sigma_1 - 1)$. Podle tohoto znaménka lze pak rozlišovat kladný a záporný diferenciální efekt. Při $\rho > 0$ (tedy v typickém případě) je kladný diferenciální efekt ekvivalentní vztahu $\sigma_2 > \sigma_1$ ("rozhození" souboru), zatímco negativní diferenciální efekt odpovídá vztahu $\sigma_2 < \sigma_1$ (tedy "semknutí", "normalizaci" souboru).

Veličina Y (přesněji její empirický protějšek, definovaný pomocí výběrových ekvivalentů ρ a μ) se hodí ke grafickým účelům, jmenovitě ke konstrukci rozptylového diagramu (scatter-plotu) Y proti X_1, který by měl (na rozdíl od grafu $X_2 - X_1$ proti X_1) názorně vypovídat o existenci či absenci diferenciálního efektu.

V citovaném článku je konečně prezentován test hypotézy H_0 proti alternativě $\sigma_1 \neq \sigma_2$ poměrem věrohodností. (Test není založen na výběrovém korelačním koeficientu mezi Y a X_1). Testová statistika

$$T = \frac{1}{2}\,(n - 2)^{1/2} \left(\frac{s_2}{s_1} - \frac{s_1}{s_2} \right) \Big/ (1 - r^2)^{1/2}\,, \qquad (4)$$

kde s_1 a s_2 jsou výběrové směrodatné odchylky 1. a 2. měření, r je výběrový korelační koeficient mezi X_1 a X_2 a n je počet pozorování, má rozdělení t s n-2 stupni volnosti.

Výpočet podle vzorce (4) lze případně obejít. Hypotéza H_0 je ekvivalentní s hypotézou H_{00}: cor $(X_2 - X_1, X_1 + X_2) = 0$. Obvyklý t-test hypotézy H_{00} založený na výběrovém korelačním koeficientu mezi $X_2 - X_1$ a $X_1 + X_2$ je s testem (4) algebraicky totožný.

Nejlepší z článku [1] nakonec: Lze se přít, zda Berrymu (a spoluautorům) patří prvenství při řešení podproblému CO. Co se týče podproblému JAK, nemusíme se takovou otázkou trápit: Test (4) publikovali (přivedeni, pravda, k hypotéze H_0 jinou motivací než my) nezávisle Pitman [5] a Morgan [3] v roce 1939!

- * -

Zamysleme se ještě jednou nad formulací hypotézy H_0. Můžeme k ní dojít i jinak než (jak se stalo výše) vylučovací metodou.

Předokládejme následující model pro veličiny X_1 a X_2:

$$\left.\begin{array}{l} X_1 = \mu + m + e_1 \quad , \\ X_2 = \mu + \Delta + (1 + \beta)\, m + e_2 \quad , \\ \mu,\ \Delta \text{ a } \beta \text{ jsou pevné (tj. ne-náhodné) parametry,} \\ m,\ e_1 \text{ a } e_2 \text{ jsou vzájemně nekorelované} \\ \text{náhodné veličiny s nulovou střední hodnotou,} \\ \operatorname{var} e_1 = \operatorname{var} e_2 \quad . \end{array}\right\} \qquad (5)$$

Měření X_1 lze rozložit na dvě složky: $\mu + m$, kterou můžeme chápat jako určitý dlouhodobý standard objektu, a e_1 - momentální odchylku od standardu. Měření X_2 můžeme rozložit na tři komponenty: $\mu+m$, tj. "starý" standard, $\Delta + \beta\, m$ - změnu standardu po zásahu, a konečně e_2 - momentální odchylku od "nového" standardu.

Chápeme-li standard $\mu + m$ jako "podstatnou" komponentu počáteční hodnoty a změnu standardu $\Delta + \beta\, m$ jako "podstatnou" komponentu změny, popř. jinak řečeno jako "skutečný" efekt, vidíme, že hypotéza H_0 neplatí, právě když (pomineme-li případ var m = 0) $\beta \neq 0$, tj. když "skutečný" efekt závisí na "podstatné" komponentě počáteční hodnoty. "Podstatná" závislost je tedy z tohoto pohledu závislostí mezi "podstatnými" komponentami měření.

Místo o "podstatných" komponentách měření bychom mohli mluvit také o míře určité vlastnosti organismu (stálejší než okamžité hodnoty měřené veličiny); pak se můžeme vrátit k "definici" třídy W - uvidíme, že pro její prvky je charakteristická záměna poznatků o momentálních hodnotách měření za poznatky o míře vlastnosti. (V kostkách: z výkonu se stává výkonnost.).

Naposledy k paviánům: Možná, že bylo zjištěno více, než jen vztah (2). "Zhruba konstantní" veličina je v překladu do jiného dialektu veličina s velmi malým rozptylem; není vyloučeno, že byl nalezen skutečný (ve smyslu práce [1]) diferenciální efekt.

Není mnoho divného na tom, že se Ian Stewart pohoršuje nad vyjádřením tak prostého faktu - že D je zhruba konstantní - složitou hatmatilkou o vztazích mezi počáteční hodnotou a efektem. Berryho práce [1] i naše úvahy o modelu (5) nicméně odvádějí podobné dílo (co asi Ian Stewart na to?): legalizují překlad tvrzení o rozptylech do řeči závislostí a naopak.

Mohli bychom tedy končit v optimistické tónině. Jen ten háček kdyby věc neměla... Logické spojení mezi tvrzeními o rozptylech na jedné straně a závislostech na druhé straně je křehké - podstatně vázané na model. Zobecníme-li model (5) vypuštěním předpokladu var e_1 = var e_2, můžeme pro porušení H_0 nalézt okamžitě jiné vysvětlení, než to jediné možné v rámci modelu (5). Toto vysvětlení může vyjadřovat zajímavý poznatek o efektu zásahu (po zásahu se mohla "zpřísnit" fyziologická regulace hodnot sledované veličiny), který nicméně se závislostí na počáteční hodnotě nemusí mít nic společného.

- * -

Podívejme se závěrem na problematiku vztahu efektu a počáteční hodnoty z jiné strany. Tím, že jsme diskvalifikovali kritérium (2) jako plnohodnotnou výpověď o účincích zásahu, zlikvidovali jsme de facto jeden dosti vydatný zdroj publikovatelných výsledků. Za to jistě nemusíme dostat od každého pochvalu.

Pokud by někomu bylo výsledků založených na vztahu (2) líto, můžeme mu doporučit naše předcházející úvahy prostě ignorovat - jistě nevejdou do obecného povědomí tak rychle, aby se v rozšiřování třídy W nedalo ještě nějaký čas pokojně pokračovat.

To však není všechno, co můžeme nabídnout. Ověřování vztahu (2) je dosti spolehlivým mechanismem generování statisticky signifikantních výsledků, přece jen však mechanismem ne zcela dokonalým. Typickým zdrojem obtíží může být přítomnost pozitivního diferenciálního efektu ("rozhození" souboru) - vztah (2) v takové situaci nemusí vůbec platit nebo může být těžké ho prokázat (dosažení statistické významnosti může vyžadovat enormní počet pozorování). Předcházející úvahy nás vyzbrojily vhledem do vztahu mezi efektem zásahu a počáteční hodnotou natolik, že nyní můžeme snadno navrhnout ještě spolehlivější generátor statistické významnosti.

Všech problémů se zbavíme tím, že závislost mezi počáteční hodnotou a její změnou budeme studovat výhradně u kategoriálních, konkrétně dvouhodnotových veličin. Hodnoty spojitých veličin kategorizujeme uměle, nejlépe rozdělením na obor fyziologických a patologických hodnot; pokud jsou hodnocené veličiny a priori binární, tím lépe.

Mějme pokus, v němž jsou hodnoceny účinky léků na pacienty, jež třídíme jednak před zásahem (léčením), jednak po něm pokaždé do tříd zdravých a nemocných. Data z pokusu lze uspořádat do čtyřpolní kontingenční tabulky - viz. tab. 1. Protože studujeme vztah mezi počátečním stavem a efektem léčení, přeuspořádáme data do tabulky 2 x 3 dle tab. 2.

Tabulka 1

		KONEC		
		zdraví	nemocní	celkem
ZAČÁTEK	zdraví	a	b	a + b
	nemocní	c	d	c + d
	celkem	a + c	b + d	N

Tabulka 2

		EFEKT			
		zlepšení	beze změny	zhoršení	celkem
ZAČÁTEK	zdraví	0	a	b	a + b
	nemocní	c	d	0	c + d
	celkem	c	a + d	b	N

Abychom mohli oprávněně tvrdit, že čím je horší výchozí stav, tím je lepší výsledek léčení, potřebujeme prokázat závislost mezi řádkami a sloupci šestipolní tabulky č. 2. Kritická hodnota χ^2-testu nezávislosti v kontingenční tabulce rozměrů 2x3 na hladině významnosti 5% je rovna 5,99 - stačí nám tedy dosáhnout hodnoty $\chi^2 = 6$. Příspěvek nulových políček (odpovídajících uzdraveným zdravým a ochořelým nemocným) do statistiky χ^2 je roven

$$(a + b)\,c / N + (c + d)\,b / N = p_z\,c + (1 - p_z)\,b \quad ,$$

kde

$$p_z = (a + b) / N.$$

Koeficient p_z je roven podílu zdravých pacientů v souboru na začátku pokusu a jeho hodnota je v rukou experimentátora. Pokud podíl zdravých a nemocných vstupujících do pokusu není vůbec řízen, je postačující podmínkou signifikance χ^2-testu, aby min (b, c) $\geq$ 6. Pokud zvolíme např. $p_z = 1/2$, tj. stejný počet zdravých a nemocných v pokusu, stačí k dosažení významného výsledku, aby b+c bylo větší nebo rovno 12, tj. aby se změnil - k lepšímu nebo horšímu - stav celkově aspoň 12 pacientů.

Právě se nám začíná rýsovat nová třída výsledků biomedicínského výzkumu analogická třídě W. Formulačně jsou výsledky z obou tříd shodné, liší se jen kritéria přijetí: vztah (2) je nahrazen porušením nezávislosti řádkového a sloupcového faktoru šestipolní tabulky (dokazovaným pomocí χ^2-testu). Novou třídu by estét mohl - s ohledem na slova "Wilder" a "Categorial" - nazvat WildCat; autor práce se však přimlouvá za stručnější a nikoli zavádějící označení WC.

To ještě není všechno. Test významnosti šestipolní tabulky č. 2 můžeme implementovat ve zvláště k tomu určeném programovém systému. Protože pokusy výše uvedeného typu budou úspěšné s naprostou pravidelností, bude možno slovní interpretaci jejich výsledků dříve či později prohlásit za přírodní zákony. Zamýšlený programový systém pak bude vlastně systémem pro automatizované generování přírodních zákonů. Vzhledem k tomu, že náš systém negeneruje jen hypotézy, jako systém GUHA, ale rovnou zákony (laws), bude vhodné pojmenovat jej GULA. A protože přijetí výsledků je vlasně předem jisté, lze skutečné pokusy nahradit myšlenkovými experimenty a systém GULA redukovat na generátor textů o objektech, zásazích a veličinách, jejichž pojmenování budou jedinými vstupními údaji.

S ohledem na vysoké aplikační ambice systém GULA i na nesporné přednosti jeho projekce a programování z ekonomického a organizačního hlediska odstupuje autor těchto řádek od další účasti na všech ostatních projektech vývoje statistického programového vybavení s přesvědčením, že tvorba příspěvků do WC je maximálním efektem, kterého lze v této oblasti v současných podmínkách dosáhnout.

P. S. 1991

1.) Z literatury k danému problému jsem dodatečně vypátral dvě přehledné práce - starší Bereiterovu [11] a relativně novou Hayesovu [12]. Hayes [12] se mj. zabývá modelem, který se od (5) liší tím, že Δ je náhodná veličina; závislostí efektu na počáteční hodnotě se pak samozřejmě rozumí $\beta \neq 0$, nikoli var $X_1 \neq$ var X_2. Jestliže jsem se divil, že se Berry [1] nezmiňuje o Wilderově zákonu, pak Hayes [12] nejenže také ne, ale necituje ani Berryho [1].

2.) Nerovnost (2) má asi s živou přírodou (včetně pelikánů) přece jen víc společného, než jsem byl ochoten připustit. Že se (bezchybné) měření veličiny X dá modelovat jako součet konstantního dlouhodobého standardu a momentální odchylky s rozdělením stabilním v čase, není tak úplně samozřejmé - hodnoty X by se mohly chovat také třeba jako náhodná procházka (resp. Wienerův proces). Nerovnost (2) jde jistě v mnoha případech měření "rozumných" veličin u živých organismů na vrub existence fyziologických regulačních mechanismů. Hrací kostka ovšem dosahuje obdobné stability v čase (a tím i nerovnosti (2)) podstatně jednodušším způsobem.

3.) Odhad statistiky χ^2 šestipolní tabulky 2 lze zlepšit:

$$\chi^2 \geq c \,.\, p_z / (1 - p_z) + b \,.\, (1 - p_z) / p_z \quad .$$

K dosažení signifikance tedy stačí, aby se při shodném počtu zdravých a nemocných v pokusu změnil stav jen 6 (místo 12, jak plynulo z původního odhadu) pacientů. Stačí ale také např., a v tom především je význam zlepšení odhadu, léčit nemocné do prvního uzdravení (c = 1) a pak doplnit soubor nemocných šestinásobným počtem zdravých. Ještě pohodlnější je počkat na první onemocnění zdravého dobrovolníka (b = 1) a soubor doplnit dostatečným počtem nemocných; léčení nemocných lze přitom věnovat úsilí úměrné významnosti počtu uzdravených z hlediska významnosti.

Tyto výsledky umožní ovládnout vědu tak přesvědčivě, že si nová verze systému GULA zaslouží za genialitu připojit na konec názvu ještě jedno G.

LITERATURA

[1] Berry, D. A. - Eaton, M. L. - Ekholm, B. P. - Fox, T. L.: Assessing Differential Drug Effect. Bi‹ etrics *40* (1984), 1109 - 1115.

[2] Jordan, H. - Reinhold, D. - Wagner, H.: Kritische Untersuchungen zur Anfangs-Endwert-Problematik von Kardiodynamischen Messgrössen unter dem Einfluss einer Bäderkur. Zeitschr. Innere Medizin *19* (1964) 897 - 901.

[3] Morgan, W. A.: A test for the significance of the difference between the two variances in a sample from a normal bivariate population. Biometrika *31* (1939), 13 - 19.

[4] Pirlet, K. - Keller, H.: Scheineffekt bei der gruppierenden grafischen Darstellung von Behandlungsergebnissen. Archiv für physikalische Therapie *20* (1968), 285 - 290.

[5] Pitman, E. J. G.: A note on normal correlation. Biometrika *31* (1939), 9 - 12.

[6] Proppe, A. - Bertram, G.: Bemerkungen zum Wilderschen Ausgangswertgesetz. Strahlentherapie *88* (1952), 573 - 596.

[7] Stewart, I.: Provozování matematiky bez povolení. Pokroky matematiky, fyziky a astronomie *29* (1984), 281 - 282.

[8] Wagner, H.: Scheinbare und echte paradoxe Reaktionen im biometrischen Modell. In: Biokybernetik, díl 1. Karl-Marx-Universität, Leipzig (1968), 282 - 287.

[9] Wilder, J.: Das "Ausgangswert-Gesetz" - ein unbeachtetes biologisches Gesetz; seine Bedeutung für Forschung und Praxis. Klin. Wochenschr. *10* (1931), 1889 - 1893.

[10] Wilder, J.: Stimuls and Response. The Law of Initial Value. Wright, Bristol 1967.

[11] Bereiter, C.: Some persisting dilemmas in the measurement of change. In: Harris, C. H. (ed.): Problems in measuring change. Univ. of Wisconsin Press, Madison 1963, 3 - 20.

[12] Hayes, R. J.: Methods for assessing whether change depends on initial value. Stat. in medicine 7 (1988), 915 - 927.