

Jan Frouz – Olga Vindušková

Čtení a psaní odborného textu v environmentálních vědách



KAROLINUM

**Čtení a psaní odborného textu
v environmentálních vědách**

**Jan Frouz
Olga Vindušková**

Recenzovali:

prof. RNDr. Bedřich Moldan, CSc., dr.h.c.

prof. RNDr. Tomáš Cajthaml, Ph.D.

Vydala Univerzita Karlova

Nakladatelství Karolinum

Redakce Václav Hozman

Obálka, grafická úprava a sazba DTP Nakladatelství Karolinum

Vydání první

© Univerzita Karlova, 2017

© Jan Frouz, Olga Vindušková, 2017

ISBN 978-80-246-3815-7

ISBN 978-80-246-3822-5 (online : pdf)



Univerzita Karlova
Nakladatelství Karolinum 2017

www.karolinum.cz
ebooks@karolinum.cz

Tuto knihu věnuji své ženě Jarce a dětem Markétě a Honzíkovi.
Jan Frouz

Tuto knihu věnuji svým babičkám, jejichž životní síla,
poctivost a činorodost jsou pro mě velkou inspirací.
Olga Vindušková

Obsah

1 Úvod	9
2 O poznání a o sdělení	11
3 Odborné publikace	19
3.1 Co je publikace	19
3.2 Druhy odborných publikací	23
3.3 Struktura hlavních typů publikací a jak je číst	29
4 Studium odborné literatury	33
4.1 Proč a jak čteme odborný text	33
4.1.1 Proč a jak čteme odborný text – příklad	36
4.1.2 Odbornou literaturu čteme na základě plánu	40
4.1.3 Odbornou literaturu čteme z nějakého důvodu	41
4.1.4 Odbornou literaturu čteme kreativně	42
4.2 Plánujeme četbu podle našeho cíle	46
4.3 Hledáme zdroje	52
4.4 Jak sehnat práce, které potřebujeme přečíst	56
4.5 Jak číst	56
5 Píšeme odborný text	59
5.1 Píšeme rešerši	59
5.1.1 Příprava textu	62
5.1.2 Psaní	67
5.1.3 Kontrola a opravy textu	71
5.2 Píšeme originální článek z původních dat	74
5.2.1 Posíláme článek a komunikujeme s redakcí recenzovaného časopisu	84
5.3 Další formy vědeckých sdělení	87
6 Slovo závěrem	90
7 Summary	91
8 Literatura	92
9 Rejstřík	93

1 Úvod

Pojďme si ukázat neobyčejné možnosti obyčejného textu na příkladu knihy *Eaters of the Dead* od Michaela Crichtona (Crichton, 1976), později vydané též jako *The 13th Warrior*, aby korespondovala s názvem filmu založeném na tomto příběhu. Část této knihy vychází z rukopisu Ahmad ibn Fadlana, který byl v 10. století našeho letopočtu vyslán chalífou Bagdádu Al-Muqtadirem jako velvyslanec ke králi Bulharů sídlících na březích dnešní řeky Volhy. Ahmad ibn Fadlan se zde spojil se skupinkou Vikingů a absolvoval s nimi dlouhou a dobrodružnou cestu patrně do oblasti dnešního Švédska. Pozoruhodné na tom je, že část knihy je převyprávěním původního rukopisu z 10. století, který nám přináší jednu z nejlepších autentických zpráv o životě Vikingů. Rukopis přitom Ahmad ibn Fadlan vypracoval na příkaz chalífy jako jakousi zprávu ze služební cesty. Máme zde po tisíci letech očitě (a při porovnání s dnešními archeologickými nálezy velmi přesné) svědectví člověka, který se potkal a žil s Vikingy. To je možné jen díky vynálezu písma, které umožňuje čtenáři, aby sdílel myšlenky nepřítomného pisatele. Tato schopnost zachytit myšlenky písmem umožňuje jejich sdílení nad propastí prostoru a času. Díky opisování textů a později s objevem knihtisku a rozvojem dalších sdělovacích prostředků umožnilo písmo šířit myšlenky i mezi mnoho čtenářů najednou. Byl to mimořádný vynález srovnatelný s rozvojem řeči jako takové. Řeči, která nám umožnila rozvíjet abstraktní myšlení, verbalizovat naše pocity, sprádat plány na lov kořisti, která ještě ani nepřišla na dohled, či poslat raketu do vesmíru.

Rozvoj písma měl mimořádný význam v řadě lidských činností. Zcela zásadní pak byl jeho význam pro rozvoj vědy. S nepatrnou nadsázkou bychom mohli říci, že rozvoj písma byl jedním ze základních předpokladů, které vůbec umožnily rozvoj vědy jako takové.

Jak jsme již řekli, písmo umožňuje přítomnému pisateli zachytit své myšlenky pro nepřítomného čtenáře. Je to obrovská vymoženost a také jedno z největších úskalí při používání písma. Hovoříte-li s někým tváří v tvář, máte kromě řeči řadu dalších informací – je tu místo a situace, kde se rozhovor odehrává, je tu řeč těla, výraz tváře. Význam řeči samotné je doplněn intonací, rytmem a tónem hlasu. Je zde velké množství informace, která je v kontextu. Tuto informaci je ale velmi těžké jednoznačně přenést do slov. Věda je navíc založena na sdílení velkého množství informací mezi lidmi různých národů, kultur a jazyků, z čehož plyne, že velká většina pisatelů i čtenářů dnešní odborné literatury komunikuje v jiném než ve svém rodném jazyce. Proto si řada disciplín vyvinula specifický styl. Jeho používání komunikaci formalizuje, možná i trochu zplošťuje, ale do jisté míry nám nahrazuje kontext. Stalo se zvykem doprovázet určitou informaci dalšími, na které by se čtenář mohl zeptat. Vysvětlit čtenáři od základu, proč jsme to celé dělali. Předvídat některé otázky čtenáře. Dopředu přinést důkazy ke svým tvrzením. Oddělit změřená fakta od svých vývodů a domněnek. Popsat postupy, které jsme pro získání

faktů použili. Jasně oddělit vlastní myšlenky od myšlenek převzatých od jiných lidí a ty navíc identifikovat tak, aby se čtenář mohl seznámit s jejich původní podobou atp. To samozřejmě celé sdělení trochu komplikuje, ale čtenář má určitou praxi v používání daného stylu a dokáže z takového sdělení rychle získat potřebné informace.

Obrovský rozvoj technologií na poli médií, šíření a zpracování informací vedl k explozi dostupných informací. Je jich tolik, že je nelze všechny prostudovat a že vybrat pouze ty správné, které bychom studovat měli, představuje samo o sobě nelehký oříšek.

Jmenovaná úskalí představují jen zlomek těch, které na poli užívání a tvorby vědeckého textu můžete potkat. Následující text si klade za cíl upozornit čtenáře na podobná úskalí a poskytnout mu základní vodítko, jak jim čelit. Text je určen studentům vysokých škol, kteří se seznamují s odbornou literaturou a sepisují své kariérní práce od bakalářské po disertační, případně se pokoušejí publikovat své výsledky ve vědeckých časopisech. Text je na řadě míst doplněn rámečky (boxy). Jejich význam je několikový: snaží se doplnit základy souvisejících disciplín nutných k porozumění odbornému textu (např. boxy o statistice) či rozvíjí a doplňují vyložený text.

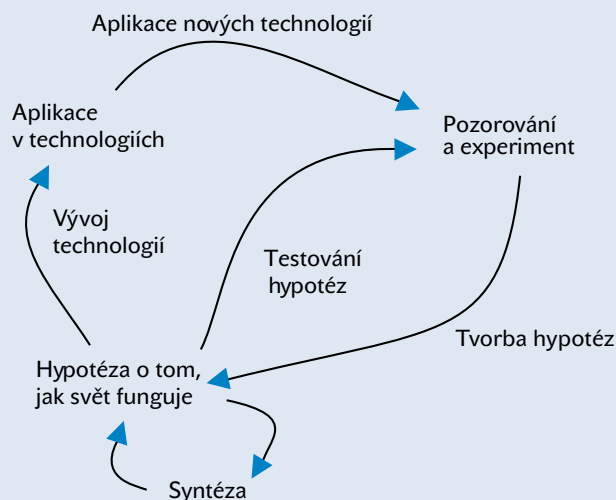
2 O poznání a o sdělení

Jistě jste již četli mnoho učebnic, které obsahovaly informace a svého druhu skládaly určitý příběh. Ale jak se vlastně ty informace do učebnic dostaly? Zde uděláme krátkou odbočku a povíme si něco o poznávacím procesu. Začneme žertem o pomateném vědci, který trochu sadisticky studoval blechu. Postavil si ji na stůl, klepl za ní prstem a přitom řekl: „Blecho, skoč!“ Blecha skočila. Učenec si to zapsal, pak utrl bleše jednu nohu a celý postup opakoval. Na slova: „Blecho, skoč!“ blecha opět skočila, učenec si to zapsal, pak utrl bleše další nohu a celý postup opakoval se stejným výsledkem. Tak to šlo dál a dál, dokud bleše neutrl poslední nohu a kdy náhle na slova: „Blecho, skoč!“ blecha zůstala ležet na místě. Učenec z toho vyvodil, že utrhnete-li bleše všechny nohy, blecha neslyší. Podíváme-li se na tuto smyšlenou příhodu ještě jednou, vidíme, že tu máme dvě věci – fakta (v našem případě chování blechy za určitých podmínek) a potom interpretaci těchto fakt, tedy více či méně opodstatněnou domněnku, která vysvětluje, proč jsou fakta taková, jaká jsou. Je důležité uvědomit si, že interpretace vychází z fakt, nikoli obráceně (Box 1). Zní to triviálně, ale řada učebnic předestře nejprve nějakou obecnou myšlenku a pak buď požaduje, abychom jí věřili, nebo ji v lepším případě demonstruje na příkladech. To samo o sobě není nic špatného, bereme-li stále na vědomí, že postup, „jak se na takovou věc přijde“, je právě opačný. Napřed máme totiž řadu jednotlivých příkladů (faktů), z nich potom dovodíme nějaký mechanismus vysvětlující, proč jsou fakta taková, jaká jsou (interpretaci), a ten pak testujeme na dalších příkladech (zobecnění interpretace). Poslední dva kroky mohou (často) přijít i obráceně – zjistíme-li, že řada věcí se chová podobně, tedy že řada faktů ukazuje podobným směrem, hledáme poté interpretaci, která je vysvětlí všechny najednou. Odvození obecného tvrzení z konkrétních faktů se v klasické logice označuje jako indukce (neboli zobecňování) a jedná se o klíčový nástroj poznání v empirických vědách. Opakem indukce je dedukce, při níž naopak z obecných předpokladů vytváříme konkrétní závěr. Dedukci využíváme například při zpětném ověřování/vyvracení obecných tvrzení a vytváření pracovních hypotéz.

Ilustrujme to ještě na jiném pokusu. Představte si, že budete topit v kamnech nebo v krbu a bude vás zajímat, proč vlastně věci hoří. Půjďte na to „vědecky“ a zvážíte si všechno palivo, které přiložíte, a následně popel, který vám vznikne. Budete pálit dřevo, papír, uhlí a ve všech těchto případech se ukáže, že popel je lehčí než původní palivo. To vás může vést k domněnce (vytvořené indukci), že spalitelné látky obsahují část, která hoří, a část, která nehoří. Což je v podstatě správná úvaha. Můžete jít ještě dál a předpokládat, že hoření je ztráta oné hořlavé části, jakési hořlavé substance, a že tato hořlavá substance je u všech hořlavých látek stejná. Tato představa byla základem takzvané flogistonové teorie hoření. Předpokládala, že všechny hořlavé látky obsahují flogiston a že hořením se flogiston ztrácí (přemění se na oheň). Nicméně později se ukázalo, že některé

1 Proces poznávání reality

Realita jako taková existuje mimo naše vědomí a v podstatě je vědomím neuchopitelná. V našem vědomí pracujeme s nějakou více či méně dokonalou představou reality. Pro empirické vědy, mezi které environmentální vědy patří, je charakteristické, že tato představa vzniká na základě pozorování učiněných pomocí smyslů, přístrojů atp. Nicméně přístroje a smysly nám neodkryjí celou realitu, poskytují pouze určitá jednotlivá fakta. Na základě těchto vjemů a dat si vytváříme nějakou představu (hypotézu) o realitě a o tom, jak svět funguje. Následně se snažíme na základě těchto představ předpovědět, jak by se věci měly chovat za určitých podmínek. Tuto představu pak testujeme pomocí experimentů. Tomuto postupu říkáme testování hypotéz. Soubor hypotéz, které se experimenty nepodařilo vyvrátit, syntetizujeme do komplexních představ o tom, jak svět funguje. Tyto představy spolu s dalšími pozorováními umožní formulovat další představy (další hypotézy), které pak testujeme pomocí dalších experimentů, a celý koloběh se opakuje. V důsledku poznání také vznikají aplikace, které zlepšují naše technologie, zpřesňují naše přístroje nebo nám umožňují měřit věci, které jsme do té doby měřit neuměli. Pozorování s využitím lepších přístrojů nám opět umožňuje formulovat nové představy o tom, jak svět funguje, nebo testovat představy, které jsme předtím z technických důvodů testovat nemohli.



kovy jako například hořčík při hoření na váze získávají, což odporuje flogistonové teorii. Vědec, který se rozhodl pálit hořčík, si nejspíše před experimentem vytvořil na základě dedukce konkrétní domněnku o tom, jak experiment dopadne. Protože 1. hořčík je hořlavý a 2. hořlavé látky ubývají při hoření na váze, domníval se, že hořčík bude hořením ubývat na váze. V momentě, kdy se jeho domněnka nepotvrdila, bylo jasné, že alespoň jeden z jeho dvou obecných předpokladů musí být nesprávný. Jiné experimenty zároveň ukázaly, že zahrneme-li i plynné produkty hoření látek, jako je dřevo, uhlí atp., pak je celková hmotnost produktů hoření větší než hmotnost původního paliva. Tato zjištění vedla

k vyvrácení flogistonové teorie, která byla nahrazena současnou představou, že podstatou hoření je oxidace látek. Na tomto příkladě je zajímavé, že pokus podporující a popírající existenci flogistonu mohl vlastně vypadat velmi podobně. V prvním případě bychom spálili na miskách vah kus papíru a shledali, že popel je lehčí než papír. V druhém případě bychom umístili papír do vzduchotěsně uzavřené nádoby naplněné kyslíkem a zažehli bychom ho třeba elektrickým výbojem. Pak bychom shledali, že váha kyslíku a paliva na začátku se rovná váze zplodin hoření na konci. Tento příklad nám ukazuje, že k tomu, abychom mohli získaná fakta správně interpretovat, musíme vědět, jakým **postupem** byla získána. Chyba v postupu, nebo prostě jiný postup, může znamenat odlišný výsledek. Například výše uvedený výsledek by nevyšel, kdyby nádoba nebyla skutečně vzduchotěsná nebo nebyla naplněna kyslíkem atp. Přesný popis metody, jak byly výsledky získány, je důležitý ještě z jednoho důvodu – umožňuje zopakování (**reprodukcí**) daných výsledků někým jiným. Jiní badatelé si mohou na základě původního popisu sestavit podobnou aparaturu a na ní se pokusit původní výsledky zopakovat. Právě tato opakovatelnost (reprodukovatelnost) pokusů a tím ověřitelnost faktů je jedním ze základních pilířů moderní vědy. Navíc s postupem našeho poznání jsme často schopni sestavit lepší a přesnější aparatury či přístroje, které nám umožní to či ono měřit. To nám umožňuje provádět pokusy přesněji a často tím získáme nová fakta a nové pohledy na realitu, které můžeme nově interpretovat.

Tím jsme si mimoděk představili, jak vzniká jeden z typů písemné odborné práce, tzv. **práce empirická**, která nás seznamuje s výsledky empirického výzkumu získanými reprodukovatelným způsobem. Taková práce obsahuje dané výsledky (fakta), jejich interpretaci a popis metody, jak byly tyto výsledky dosaženy. Navíc obsahuje také nějaký úvod, který seznamuje čtenáře s tím, proč bylo důležité daná fakta zjistit, jak doplní to, co o problému už víme atp. A tady se nám ukazuje důležitá spojnice mezi naší představou a tím, jak pokus provádíme. Jde o to, že ve výše uvedeném případě jsme měli podezření, že s flogistonem není něco v pořádku. Za prvé jsme si mohli všimnout, že řada hořlavých látek bez přístupu vzduchu nebo kyslíku nehoří, a to nehraje – flogiston by nic takového nepotřeboval. Na druhou stranu zastánci flogistonu by mohli říci, že přístup vzduchu či kyslíku je jen jedna z podmínek pro to, aby se flogiston mohl hořením uvolňovat. Mezi další podmínky by jmenovali teplo, neboť víme, že látku je třeba zahřát na určitou teplotu, aby začala hořet. Máme tedy dvě alternativy, buď jsou kyslík a teplo jen podmínky nutné pro to, aby se z látky uvolnil flogiston, který se přemění v oheň, anebo je hoření formou oxidace, při níž se uvolňuje energie na úkor chemických vazeb v hořící látce. V prvním případě nám něco (flogiston) ubude, v druhém nikoliv. Tím se dostáváme k dalšímu zajímavému rysu empirických prací, tedy prací přinášejících jednotlivá fakta, a sice že jsou často uspořádány tak, aby daly různý výsledek, platí-li jedna ze dvou alternativ interpretace předchozích pokusů. Shrňme-li si to, máme nějaká fakta získaná pomocí dané metody a z nich vyvodíme nějakou představu o tom, jak svět v dané věci funguje. Někdy můžeme takových představ mít více.

Pak můžeme uspořádat okolnosti tak, aby se výsledky dalšího měření (dalšího pokusu) musely lišit, platí-li jedno z alternativních vysvětlení. Takovému vysvětlení říkáme **hypotéza** – platí-li nějaký předpoklad, pak se realita bude chovat nějak, neplatí-li, bude se chovat nějak jinak. Často tedy nesbíráme jednotlivá fakta nazdařbůh, ale proto, abychom podpořili nebo vyvrátili naše hypotézy založené na naší dosavadní představě o světě. Někdy však nezačínáme experimentem, ale pozorováním – sledujeme, jak se věci v přírodě za určitých podmínek chovají, a až z toho vyvodíme nějakou představu o tom, jak svět v dané věci funguje. Tuto představu pak konfrontujeme s dalšími experimenty nebo s výsledky jiných pozorování a zbytek je již podobný.

Studenti i začínající badatelé často mají hypotézy za cosi složitého, co neumí dobře uchopit. Zkusme si tedy představit něco přízemnějšího. Řekněme, že jsme ztratili peněženku. Můžeme jistě začít pobíhat zcela chaoticky sem a tam (často i po místech, kde jsme před tím vůbec nebyli) a dívat se, zda peněženku neuvidíme. Jinou možností je, že si vybavíme, kde jsme se ztracenou věcí naposledy manipulovali, a myšlenkově si zrekonstruujeme cestu, kterou jsme od té doby ušli, půjdeme po ní zpět a budeme danou věc hledat. Můžeme jít myšlenkově ještě dál a přemýšlet, co jsme po cestě dělali a zda nějaká ta činnost nezvyšuje pravděpodobnost, že jsme tam mohli peněženku ztratit (třeba jsme tahali z kapsy kapesník a peněženka mohla přitom vypadnout). V každém případě máme soubor nějakých logických předpokladů o tom, jak se věci staly nebo mohly stát, a podle nich budeme řídit svoje hledání. Hypotézy jsou také takové představy o tom, jak věci vypadají, jak fungují, proč se to či ono děje atp., které si učiníme na základě toho, co o věcech víme.

Studenti si také častou pletou hypotézy s výzkumnými **otázkami**. V našem příkladu s peněženkou by výzkumnou otázkou bylo: „Kde jsem ztratil peněženku?“ Výzkumné otázky jsou často obecnějšího charakteru a jsou zakončeny otazníkem. K jejich zodpovězení musíme přijít s jednou či více hypotézami, které jsou již konkrétněji formulovanými představami, jejichž platnost můžeme ověřit. Hypotéza je tvrzení a je větou oznamovací. Hypotézou v našem příkladě by mohlo být: „Peněženku jsem zapomněl doma v kapse u kabátu.“ Tuto hypotézu bychom již mohli ověřit tím, že bychom se vrátili domů a do kabátu se podívali.

Experimenty a pozorování jsou často komplikovány tím, že věci se ne vždy chovají stejně. Co říci o pokusu, který 17× vyšel, jak jsme čekali podle hypotézy 1, a 3× vyšel podle zcela protichůdné hypotézy 2? Můžeme říci, že platí hypotéza 1? Naše přístroje také neměří úplně přesně. Co máme například říci o výsledku pokusu, když místo očekávané váhy 500 gramů jsme dostali 5 čísel, z nichž ani jedno není 500, ale všechny jsou „blízko“ – třeba 501, 498, 499, 502 a 501? Znamená to, že jsme se nikdy netrefili, anebo že jsme se vždy trefili „docela blízko“? K řešení takovýchto a podobných otázek nám slouží **statistika** (Box 2).

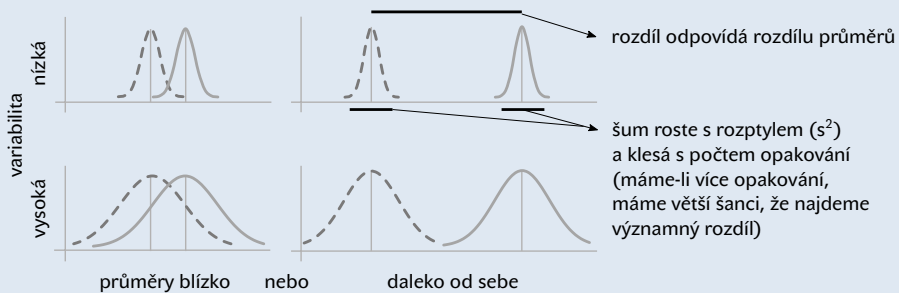
Tohle všechno ovšem jen těžko může být (a většinou není) dílem jednoho člověka. Aby celý proces poznávání reality mohl fungovat – od popisu jednotlivých jevů, přes

jejich interpretaci, po formulaci nových hypotéz, plánování, provádění a vyhodnocení dalších pokusů, jejich porovnání s dřívějšími výsledky a další syntézu – musí větší množství lidí sdílet informace a své myšlenky. K tomu slouží **odborná literatura**. Ta

2 Testování hypotéz a statistika

V tomto boxu zveme čtenáře na krátkou exkurzi (či spíše skok) do světa statistiky, která je nepostradatelným nástrojem pro interpretaci vědeckých dat. Zájemce o hlubší pochopení a ovládnutí statistických nástrojů odkazujeme na specializovanou literaturu věnovanou statistice v přírodních vědách (Zichová, 2008; Zvára, 2013; Lepš a Šmilauer, 2016; Zvárová, 2016). Zmínili jsme se o tom, že experimentálně můžeme nastavit podmínky tak, aby se příroda musela zachovat jedním způsobem, pakliže platí nějaký předpoklad (syn. hypotéza), a jiným způsobem, pokud tento předpoklad neplatí. Problémem je, že chování zkoumaného jevu mohou často ovlivňovat ještě rušivé vlivy, a tak se může stát, že se příroda nezachová vždy stejně. Řekněme, že se podle předpokladu příroda zachová často, ale ne vždy. Jenže problémem je, že nezachová-li se tak vždy, ale jen často, mohou vám vaši oponenti říci, že je to náhoda. K tomu abychom mohli říci, co je náhoda a co náhoda není (resp. co se již natolik odchyluje od náhody, že je velmi nepravděpodobné, aby to náhoda byla), nám slouží obor statistiky, který se jmenuje **testování hypotéz**.

Pojďme si tento koncept vysvětlit na příkladu. Jednoduchý typ pokusu vypadá tak, že máme dvě skupiny pokusných entit, např. soubory luk, a na jednom z nich provedeme nějaký zásah, např. hnojíme dusíkem nebo kosíme, a na druhém z nich takový zásah neprovedeme. Poté na obou změříme nějaké parametry, které chceme sledovat, např. počet druhů nebo celkovou biomasu. Budeme mít nějakou hypotézu, např. že hnojení dusíkem zvyšuje celkovou biomasu rostlin. Průměry obou skupin se budou skoro určitě lišit, ale ani čísla v rámci jedné skupiny nebudou zcela stejná. To, že se průměry o trochu liší, nemusí nic znamenat, může to být náhoda. Hodíte-li si třikrát tou samou hrací kostkou, zaznamenáte si jednotlivé hody a pak uděláte ještě jednou to samé, bude se průměrná hodnota hodů získaných v každé trojici hodů dost možná také trochu lišit. Vrátime-li se k našemu původnímu příkladu s loukami, pak bude-li rozdíl mezi průměry obou souborů malý a variabilita uvnitř souborů velká, usoudíme, že se soubory neliší a jedná se spíše o náhodnou variabilitu. Na druhou stranu bude-li variabilita uvnitř souborů malá a rozdíl mezi průměry velký, potom se nejspíš o náhodu nejedná a oba soubory se liší. Jedním ze způsobů, jak porovnat velikost rozdílu mezi dvěma soubory s variabilitou souborů, je dvouvýběrový **t-test** (syn. **Studentův test**). Jako míra rozdílu souborů je použit rozdíl mezi průměry jednotlivých souborů (*signal*). Složitější je odhad rušivých prvků neboli šumu (*noise*). Ten se zvětšuje s variancí jednotlivých souborů a snižuje s narůstajícím počtem opakování. Varianci (rozptyl) pak spočítáme z odchylek jednotlivých hodnot daného souboru od průměru souboru. Protože odchylky jsou kladné i záporné, nesčítáme odchylky, ale jejich druhé mocniny (čtverce). Ty jsou vždy kladné, a navíc dává tento postup větší váhu větším odchylkám. Vlastní výpočet t-testu dnes poskytují běžné tabulkové editory, např. Excel. T-test nám poskytne tzv. **p-hodnotu**, čili pravděpodobnost, s jakou se oba soubory neliší (= jsou shodné). Můžete si ji představit i tak, že jde o počet případů, ve kterých se Gaussovy křivky obou sledovaných souborů překrývají. Je-li tato hodnota malá – tradičně se tak uvažuje o hodnotách menších než 0,05 (čili pravděpodobnost shody 5 %) – pak usuzujeme, že pravděpodobnost, že se soubory neliší, je moc malá na to, abychom věřili, že se neliší. Formálně pak výsledek testu komentujeme tak, že se soubory statisticky významně liší na hladině významnosti $\alpha = 0,05$ (nebo na 5% hladině významnosti).



$$T = \frac{\text{rozdíl}}{\text{šum}} = \frac{\bar{y}_1 - \bar{y}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$\bar{x} = \frac{1}{n} \sum x_i$$

$$s^2 = \frac{1}{n - 1} \sum (x_i - \bar{x})^2$$

Vysvětlivky:

\bar{y}_1, \bar{y}_2 ... průměr souboru 1, resp. 2

s_p ... sdružená směrodatná odchylka (angl. pooled standard deviation)

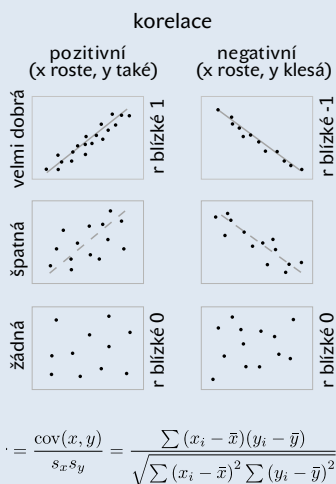
s_1, s_2 ... směrodatná odchylka souboru 1, resp. 2


n_1, n_2 ... počet opakování v souboru 1, resp. 2

T-test je jedním z nejjednodušších, ale také nejčastěji používaných statistických testů. Navíc si na něm můžeme vysvětlit logiku i dalších statistických testů. Těch existuje celá řada, ale fungují na podobném principu. Na začátku každého testování hypotézy je tzv. nulová **hypotéza** (H_0), jejíž platnost testujeme. H_0 je pro daný typ testu daná a např. pro dvouvýběrový t-test H_0 zní: Průměry dvou souborů se neliší. Na nás je zformulování tzv. **alternativní hypotézy** H_1 . Ta může být buď oboustranná (H_1 : Průměry se liší.), nebo jednostranná (např. H_1 : Průměr souboru 1 je větší než průměr souboru 2). V testu tedy testujeme platnost nulové hypotézy oproti alternativě. Nulová hypotéza a alternativa se vzájemně vylučují, takže zamítneme-li testem nulovou hypotézu, potvrdíme tím alternativní hypotézu. Při testu samotném je následně na základě našich dat vypočteno tzv. **testovací kritérium** (u t-testu se kritérium nazývá „T“ a odtud i jeho název). To je poté porovnáno s teoretickým rozdělením kritéria (u t-testu je to Studentovo t-rozdělení), které závisí na počtu opakování (u dvouvýběrového t-testu na velikosti našich dvou souborů) a představuje rozdělení hodnot T za platnosti nulové hypotézy. Z tohoto porovnání vzejde již zmíněná p-hodnota, což je vlastně pravděpodobnost, se kterou bychom za platnosti nulové hypotézy dostali z našich dat právě tuto konkrétní hodnotu testového kritéria. Pokud je tato pravděpodobnost velmi malá (resp. je menší než hranice α , kterou jsme si dopředu zvolili), zamítáme nulovou hypotézu ve prospěch alternativní hypotézy. Důležité je, že statistické testování hypotéz nám umožňuje udělat si představu o pravděpodobnosti, s jakou naše hypotéza o daném jevu (např. rozdílu souborů) platí. Není to tedy test ve smyslu pravda/lež, ale říká nám, jaká je pravděpodobnost, že jsou pozorovaná data pouze dílem náhody.


Vysvětlíme si ještě jeden významný koncept a tím je **korelace**. Korelace nám říká, jak těsně na sobě závisí dvě veličiny, nebo jinak: do jaké míry platí, že stoupne-li hodnota jedné veličiny, pak stoupne (nebo klesne) hodnota druhé veličiny a naopak. Na obrázku níže jsou uvedeny

příklady velmi dobré, špatné nebo žádné korelace. Ta může být buď pozitivní, nebo negativní. Pozitivní korelace znamená, že s nárůstem x roste i y ; při negativní korelaci bude naopak s nárůstem x hodnota y klesat. Míru těsnosti, s jakou se tak děje, měříme **korelačním koeficientem r** . Postup výpočtu je uveden níže, ale nebudeme si ho vysvětlovat. Korelační koeficient může nabývat hodnot od -1 (těsná negativní korelace; všechny body budou na přímce, která klesá) do 1 (těsná pozitivní korelace; všechny body budou na přímce, která stoupá). Je třeba si ale uvědomit, že tato těsnost závislosti sama o sobě neříká nic o tom, zda závislost je nebo není náhoda. Máte-li totiž dva náhodné body xy , budete vždy schopni jimi proložit přímku. Naopak když se mnoho bodů xy uspořádá v jedné linii, asi to náhoda nebude. K tomu bychom tedy mohli posoudit, s jakou pravděpodobností byl daný korelační koeficient dosažen náhodou, potřebujeme vědět, kolik jsme měli bodů. Obrázek níže ukazuje závislost mezi počtem opakování a hodnotou korelačního koeficientu, který bychom mohli považovat za náhodný jen s malou pravděpodobností (5% pro 0,05 nebo 1% pro 0,01). U takto malých pravděpodobností máme tendenci věřit tomu, že závislost mezi body není náhodná. Zde jen malou odbočku. To, že závislost není náhodná, neznamená automaticky, že mezi danými parametry x a y je kauzální vztah (tedy že jeden parametr určuje druhý). Existuje například velmi těsná a nenáhodná korelace mezi počtem čápů v jednotlivých okresech ČR a počtem narozených dětí v těchto okresech. To ale není důkaz toho, že čápi nosí děti. Spíše to znamená, že existuje faktor nebo faktory, které ovlivňují čápy i porodnost podobným způsobem (ve větších okresech bude více dětí i čápů, čápi mají raději nížiny, kde jsou okresy lidnatější, a proto se tam rodí více dětí atp.).

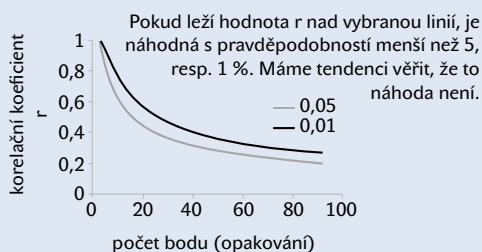




Vezmeme-li dva jakékoliv náhodné body, můžeme jimi vždy proložit přímku (r bude 1 nebo -1), ale pořád to bude náhoda.



Naopak v případě, že budeme mít hodně bodů a r vyjde blízké 1, náhoda to asi nebude.



Vysvětlivky:

x_i, y_i, \dots i -tá hodnota veličiny x , resp. y

$\text{cov}(x, y)$... kovariance veličin x a y

s_x, s_y ... směrodatná odchylka veličiny x , resp. y

Jak jsme si ukázali na příkladu s čápy, statistika je pouze pomůckou pro interpretaci dat, nikoliv zárukou její smysluplnosti. Naše zjištění se také odvíjí od přesnosti a reprezentativnosti našich dat. Ta závisí jednak na **přesnosti** našeho měření a jednak na **reprezentativnosti** vzorku, který

máme. Řekněme, že bychom chtěli odhadnout hmotnost úrody na poli s pšenicí a že si můžeme vybrat váhy, které na to použijeme, a celé měření pak těmto vahám uzpůsobit. Řekněme, že na výběr jsou laboratorní váhy, které váží až 200 g s přesností na desetitisícinu gramu, a dále váhy mostní, které váží 10 t s přesností na 5 kg. Které váhy si v takovém případě vybereme? Laboratorní váhy jsou jednoznačně citlivější a dle poměru nejmenšího možného dílku a váživosti jsou i přesnější. Přesto by bylo lepší si vybrat váhy mostní. Proč? Protože váhy mostní nám umožní zvážit úrodu z celého pole během jednoho nebo několika málo vážení. Pokud zvážíme celé pole najednou, nevnášíme do našeho měření žádnou další chybu, možná vyjma ztráty sklizených zrn. Budeme-li pracovat s laboratorními vahami, budeme muset odebírat vzorky na nějaké definované ploše třeba na 1 m². Jednotlivé metrové plošky na poli se přitom od sebe výrazně liší, často mnohonásobně, a tak budeme muset odebrat plošek vysoký počet (mnoho desítek), abychom odhadli průměr reprezentující celé pole. Navíc u každé plošky budeme řešit problém okrajů, jak zahrnout třeba spadlé klasy, dále budeme muset odhadnout plochu celého pole a to vše je zdrojem chyb. Tyto nejistoty jsou součástí každého měření, jsou jimi zasažena nejen vaše data, ale i data, která převezmete z literatury nebo ze statistické ročenky, a to i tehdy, kdy tyto nepřesnosti nejsou kvantifikovány. S nejistotami je tedy třeba v našich úvahách a závěrech počítat. S nejistotami je dobré počítat i při prezentaci dat. To nejmenší, co můžeme udělat, je, že při uvádění kvantitativní hodnoty bude přesnost uváděných čísel odpovídat přesnosti měření. Vážíme-li s přesností na kg, je nesmyslné uvádět průměrnou váhu s přesností na mg.

umožňuje sdílení myšlenek a faktů mezi lidmi, kteří se často tváří v tvář nikdy nepotkali. Rozlišujeme již zmíněné empirické (neboli **primární**) práce popisující primární data (fakta, konkrétní výsledky jednotlivých měření) a potom tzv. **sekundární** literaturu, která se zabývá následným srovnáváním, interpretací a zobecňováním těchto fakt. Patří sem různé přehledové práce, rešerše, teoretické práce popisující nové komplexní hypotézy atp. Dostáváme tedy gradient prací od empirických studií přinášejících detailní poznatky o tom, jak se chová malá část reality, po práce, které tyto poznatky více a více zobecňují. Tato zobecnění jsou samozřejmě, jak jsme si ukázali výše, podrobována testu dalšími empirickými studiemi, a tak se naše poznání posouvá v jakési smyčce pořád dopředu (Box 1). Z boxu 1 nepřímou vyplývá, že proces poznávání je kolektivním dílem, ve kterém vaše práce navazuje na práci jiných a napomáhá práci dalších.

3 Odborné publikace

3.1 Co je publikace

Publikace neboli zveřejnění znamená poskytnutí myšlenek nebo faktů ostatním, aby se jimi mohli zabývat. Ostatně slovo publikace pochází od slova *public* čili veřejný, veřejnost atp. Slovem publikace můžeme označovat jak proces uvedení nějaké informace či zprávy ve známost (nejčastěji v tisku, ale dnes stále častěji i elektronicky) – a je tedy synonymem pro slovo **zveřejnění** či vydání –, tak i **výsledek** tohoto procesu, tedy výtisk knihy nebo časopisu (před vynalezením knihtisku i opis a dnes i e-knihu nebo článek v elektronickém časopise). Takovým výstupem je **záznam textu vzniklý v mnoha kopiích, které jsou veřejně dostupné** (můžete si je koupit, získat je darem nebo vyměnit za něco jiného). Pro naše účely vystačíme s touto definicí, nicméně řada vědních oborů například zoologická nebo botanická taxonomie mají daleko specifitější a přesnější pravidla pro to, co je a není publikace. Taková definice má své praktické důsledky – například to, že autorem popisu druhu je ten, kdo popis prvně publikoval. Rokem popisu je pak rok, kdy práce vyšla (resp. vešla do oběhu). Například pokud by došlo k tomu, že nějaký časopis byl vytištěn, ale v důsledku války nebyl distribuován a dostal se do oběhu až s několikaletým zpožděním, tak by až toto datum bylo považováno za rok popisu druhů popsaných v tomto časopise. To je jen na okraj pro ilustraci toho, jaké důsledky mohou mít na první pohled nepodstatné detaily. V tomto textu se budeme zabývat publikacemi v širokém smyslu slova, a dokonce i pracemi, které vznikly jen v několika málo kopiích, jako jsou práce bakalářské nebo diplomové.

Po formální stránce jsou nejčastějšími formami publikace **časopisy**, které poznáte podle toho, že mají ISSN (International Standard Serial Number čili mezinárodní standardní číslo seriálové publikace), nebo **knihy a sborníky**, které mají ISBN (International Standard Book Number neboli mezinárodní standardní číslo knihy). ISSN a ISBN jsou číselné indikátory, které jednoznačně určují dané periodikum nebo knihu. ISBN u nás přiděluje Národní agentura ISBN fungující při Národní knihovně ČR. ISSN pak uděluje České národní středisko ISSN při Národní technické knihovně. **Seriálové publikace** jsou publikace vydávané po částech, obvykle s číselným nebo chronologickým označením, určené k neomezenému pokračování (bez ohledu na periodicitu). Některé knihy, které jsou součástí série, mají dokonce přiděleny oba identifikátory. Na rozdíl od ISBN však ISSN nemá žádnou vnitřní strukturu a nelze z něj zjistit žádné informace o publikaci – jde o jednoduchý identifikátor (z ISBN lze naproti tomu určit např. vydavatele). Některé databáze mají vlastní kritéria na periodicitu, počet čísel za rok atp., aby mohla být seriálová publikace považována za časopis.

Z číselných indikátorů publikací zmíníme ještě **DOI** (Digital Object Identifier). DOI je kód obsahující písmena a čísla, který umožňuje dohledat na internetu jakýkoli digitální dokument, který je tímto identifikátorem opatřen. Velká světová nakladatelství dnes vydávají knihy či články v časopisech nejen na papíře, ale také je v digitální podobě šíří na internetu. Často se dokonce taková díla objeví na internetu dříve, než jsou vytištěna, a existují i časopisy, které vychází jen elektronicky a na papír se již vůbec netisknou. Všechny tyto dokumenty mohou být (a dnes zpravidla i jsou) označeny unikátním DOI. DOI ale mohou nést i jiné digitální dokumenty, které by nezapadaly do výše uvedeného schématu kniha, kapitola v knize, článek ve sborníku, článek v časopise atp. DOI nám umožňuje dohledat text na internetu a obsahuje i informaci o organizaci, která DOI vydala.

Zveřejňování vědeckých výstupů nebylo vždy samozřejmé. Staré alchymistické školy tajily své postupy a ještě v počátcích moderní vědy byl často zveřejňován jen výsledek bez úplné cesty, jak se k němu došlo, a tedy bez možnosti následné kontroly. Dnes je naopak snahou co největší průhlednost a kontrolovatelnost našich vývodů. Představy o světě jsou založeny na faktech, která jsou získána za známých podmínek a tím jsou alespoň do určité míry ověřitelná. Novověk také přinesl myšlenku autorství, autorského práva a s tím spojené problémy s jeho porušením (Box 3).

3 Autor a autorská práva, plagiát a plagiátorství

Koncept autorství je historicky poměrně nový a objevuje se prvně okolo roku 1700. V našem autorském zákoně (121/2000 Sb.) je **autor** chápán jako fyzická osoba, která tvůrčím způsobem vytvořila dílo. **Dílo** je pak definováno jako jakkoli vnímatelné dílo, které je výsledkem jedinečné tvůrčí činnosti autora (121/2000 Sb.). Důraz je tedy kladen na jedinečnou tvůrčí činnost autora, která přináší novou hodnotu (u vědeckého díla tedy např. nová fakta nebo nový pohled na známá fakta). Vydávání cizího díla nebo jeho části za vlastní označujeme jako **plagiátorství**. Plagiátorství je obecně chápáno jako podvodné jednání. Lze to připodobnit k situaci, kdy se budete vydávat za vlastníka věci, která vám nepatří, s cílem získat z dané situace prospěch, například pronajmete někomu cizí byt atp. Plagiátor se vystavuje riziku právního postihu za své jednání. Tento postih může přijít v zásadě ze dvou stran – ze strany autora nebo vlastníka autorských práv, který se s vámi může soudit za to, že jste porušil jeho autorská práva, a v případě vědeckého díla i ze strany akademické obce. Vzhledem k tomu, že akademická obec je svou podstatou založena na sdílení myšlenek, je obecně plagiátorství považováno za hrubý přečin proti pravidlům komunity. Na řadě akademických pracovišť může prokázaný případ plagiátorství vést ve svých důsledcích až k propuštění (u akademických pracovníků) nebo vyloučení ze studia (v případě studentů).

V případě vědecké práce je nejčastějším případem plagiátorství použití části textu jiného díla bez toho, aby byl zdroj řádně citován. Plagiátorství se tedy můžeme dopustit i nedbalou, neúplnou nebo nepřesnou citací nebo jejím opomenutím, a to i z nedbalosti. Přitom je třeba zdůraznit, že řádně citovat je třeba i tzv. volné zdroje například z internetu.

S rozvojem internetu narůstá dostupnost textů prakticky o čemkoli, a to může svádět k jejich využití při přípravě textů nových často i metodou „copy-paste“. To na druhou stranu vedlo