

Úvod do navrhování a analýzy průmyslových experimentů

Gejza Dohnal



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Věnováno Aleši Linkovi

Obsah

1	Základy navrhování experimentů	1
1.1	Historie řízení kvality	1
1.2	Historie DOE	2
1.3	Základní pojmy	3
1.4	Návrh a analýza experimentu v pěti krocích	5
1.4.1	Analýza procesu	5
1.4.2	Návrh experimentu	6
1.4.3	Provedení zkoušek	9
1.4.4	Analýza výsledků	10
1.4.5	Závěry	11
2	Stochastické repetitorium . . .	13
2.1	Náhodná veličina	14
2.1.1	Rozdělení náhodné veličiny	14
2.1.2	Charakteristiky náhodných veličin	16
2.1.3	Normální rozdělení pravděpodobnosti $N(\mu, \sigma^2)$	19
2.2	Testování statistických hypotéz	22
2.2.1	Test statistické hypotézy	22
2.2.2	Testy o parametrech normálního rozdělení	24
2.2.3	Testy dobré shody	28
2.2.4	Některé neparametrické testy	30
2.3	Regresní analýza	31
2.3.1	Regresní závislost	31
2.3.2	Jednoduchá přímková regrese	33
3	Jednofaktorové experimenty	37
3.1	Dvouúrovňové jednofaktorové experimenty	38
3.1.1	Nezávislá měření	39

3.1.2	Závislá pozorování	42
3.1.3	Uspořádání do bloků	44
3.2	Jednofaktorové experimenty s více úrovněmi	45
3.2.1	Ordinální odezva	45
3.2.2	Nominální odezva	47
3.2.3	Latinské čtverce	48
3.3	Regresní experimenty	48
3.3.1	Jednoduchá přímková regrese	50
4	Vícefaktorové experimenty	53
4.1	Dvoufaktorové experimenty	54
4.2	Vyhodnocení vícefaktoriálních experimentů	57
4.3	Vícefaktoriální experimenty	58
4.4	Směšování ve vícefaktoriálních návrzích	58
4.5	Významné body experimentu	61
5	Taguchiho robustní návrhy	63
5.1	Designové matice	65
5.2	Řízené a šumové faktory	66
6	DOE v MINITABu	69
6.1	Program MINITAB	69
6.2	Vytvoření návrhu experimentu	71
6.3	Vstup dat	72
6.4	Vyhodnocení výsledku experimentu	74
6.5	Řešený příklad z praxe	76
6.5.1	Návrh experimentu	76
6.5.2	Příprava vzorků a měření	79
6.5.3	Analýza výsledků	80
6.5.4	Závěr	87

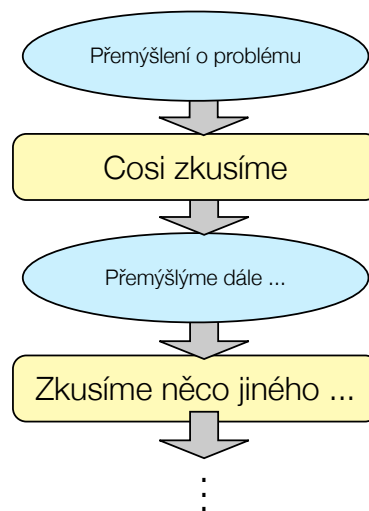
Úvod

Většina našich znalostí o okolním světě je získávána na základě experimentování. Malé děti experimentují s věcmi kolem sebe, s chováním lidí, zvířat, zkoušejí co se stane když ... Jako dospělí používáme experiment k ověření našich představ, teorií a logických úvah. K tomu využíváme znalostí a zkušeností, získaných také na základě výsledků experimentování ostatních lidí. Experimentální ověřování předpokladů a rozvíjených teorií je jednou ze základních činností ve vědeckém zkoumání. Experiment je často jediným zdrojem informace v situaci, kdy teoretické znalosti nestačí nebo je nelze použít vzhledem ke složitým reálným podmínkám.

Experiment může probíhat různým způsobem. Zřejmě nejstarším a nejčastěji používaným přístupem je takzvaná *metoda pokusu a omylu*. Tato metoda je výsledkem spontánního chování a je nejméně efektivní. Může být úspěšná pouze tehdy, je-li spojena se zkušeností a intuicí. Lze ji použít pouze v jednotlivých případech a v omezeném okruhu aplikací, přičemž výsledek je nejistý a může se dostavit až za velmi dlouhou dobu.

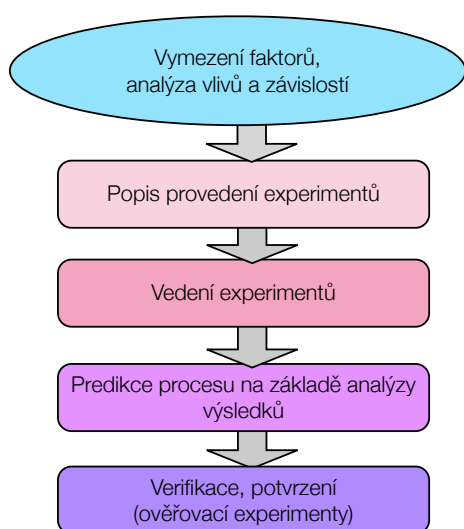
Se stále modernějšími technologickými postupy se stávají i výrobky a procesy stále více komplikovanější. Vzhledem k tomu, že náklady na experimentování se prudce zvyšují, pro analytika, který má omezené finanční zdroje a čas, se stává nemožné, aby prozkoumal množství faktorů, které ovlivňují tyto složité procesy pomocí metody pokusu a omylu.

Proto je třeba využívat techniky, které identifikují vliv důležitých faktorů tím nejefektivnějším způsobem a řídí proces nejlepšího nastavení těchto důležitých faktorů pro uspokojení stále rostoucí poptávky po zlepšení kvality a zvýšení produktivity. Techniky *plánování experimentů* poskytují výkonné a účinné metody k dosažení uvedených cílů.



Obrázek 1: Postup metodou pokusu a omylu.

Systematický přístup k řešení problémů předpokládá týmovou práci a objektivní plánování jednotlivých kroků. Tento text se zaměřuje na navrhování experimentů především ve strojírenství a technickém výzkumu. Experiment budeme chápat jako sérii zkoušek prováděných systematickým způsobem za účelem lepšího porozumění stávajícím procesům, nebo k prozkoumání nového produktu či procesu. Nástroj pro vytváření takového způsobu experimentování se označuje zkratkou DOE z anglického "Design Of Experiments". Vyhodnocení výsledků zkoušek se provádí pomocí matematicko-statistických metod. Proto se DOE často považuje za jednu z oblastí aplikované statistiky. Experimenty jsou velmi často důležitým nástrojem při řízení kvality a jsou jednou z hlavních činností při zlepšování procesů v souvislosti s aplikací zásad Six Sigma. Z tohoto důvodu jsou i součástí managementu kvality.



Obrázek 2: Systematické kroky při návrhu experimentu.

Návrh experimentu se snaží o získání maximálního množství informací s minimem vynaložených zdrojů. Pro úspěšný návrh a vyhodnocení experimentu je nutná spolupráce jak expertů z výroby, tak i techniků designérů a odborníků z oblasti aplikované matematické statistiky. Pro úspěšnou realizaci experimentu je nezbytná i podpora vrcholového vedení firmy. Ačkoli podstata DOE spočívá v efektivním vyhodnocení poměrně malého množství zkoušek, přesto může být provedení celého experimentu velmi nákladnou záležitostí. To je také – spolu s obavou o zhoršení kvality produkce po dobu experimentování – jedním z nejčastějších důvodů odmítnutí těchto metod v praktickém provozu. Proto zde hraje klíčovou roli rozhodnutí odpovědných manažerů

podniku, kteří na základě ekonomické rozvahy nařídí potřebné kroky.

Z hlediska řízení kvality je navrhování a vyhodnocování experimentů „nikdy nekončící příběh“. Výsledky předchozích experimentů jsou často vstupními informacemi pro další, navazující experimenty, které zpřesňují modely závislostí mezi odezvou a faktory, zpřesňují počet působících faktorů a jejich sledovaných úrovní. Je důležité mít tento proces pod kontrolou, neboť jen tak lze zajistit jeho efektivitu a dosáhnout požadovaných úspor při výrobě. Znalost základních zásad při navrhování a vyhodnocování experimentů patří k základním dovednostem a je důležitá jak pro ty, kteří experimenty sami provádějí, tak i pro ty, kteří využívají jejich výsledků, tedy i pro pracovníky ve vedení firmy.

Provádění a vyhodnocení průmyslových experimentů je dnes už nepředstavitelné bez softwarové podpory. Existuje řada programových produktů, které usnadní návrh a umožní analýzu výsledků experimentu. Nicméně i zde platí pravidlo, že tyto relativně mocné nástroje musí řídit odborník, který ví co a jak správně v určité situaci použít. Pokud tyto nástroje používá laik, může dojít k chybnému vyhodnocení předpokladů, výběru nevhodné metody a v neposlední řadě k chybné interpretaci výsledků.

Tento učební text si klade za cíl seznámit čtenáře se základními principy navrhování experimentů a vyhodnocení jejich výsledků. Pro snazší porozumění se předpokládá alespoň částečná znalost základních statistických pojmů, jako jsou: náhodná veličina, stochastická nezávislost, rozdělení pravděpodobnosti, distribuční funkce, hustota, střední hodnota, rozptyl a další. Proto je v první kapitole zařazeno „Statistické repetitorium“. Čtenáři, kteří tyto metody znají, mohou klidně přeskočit.

Příklady v textu uvedené jsou zpravidla zjednodušené. Důvodem je snaha o jejich lepší a názornější výklad. Skutečné reálné úlohy bývají často mnohem složitější a jejich řešení přesahuje rámec tohoto textu. Pro řešení některých příkladů jsme použili statistický program MINITAB, který obsahuje řadu nástrojů určených přímo pro DOE. Kromě tohoto programu existuje i řada dalších, více či méně složitých softwarových produktů, umožňujících v podstatě totéž. Nejjednodušším nástrojem pro vyhodnocení výsledků experimentu může být i tabulkový procesor typu MS Excel, který také v některých případech použijeme.

Kapitola 1

Základy navrhování experimentů

1.1 Historie řízení kvality

Stojíme-li před úlohou provést výběr z několika stejných či podobných výrobků které se nabízejí za stejnou či podobnou cenu, rozhodneme se na základě jejich kvality. Je tedy přirozené, že s rozvojem obchodu a hromadné výroby se stále častěji objevuje pojem *kvalita*. Obsah tohoto pojmu je přitom velmi široký a silně závisí na individualitě zákazníka. S narůstajícím množstvím nabízeného zboží však požadavky na jeho kvalitu dostávají stále více objektivní charakter. V dnešním pojetí je kvalita výrobku charakterizována souborem objektivně měřitelných znaků a je definována v normě. To umožňuje její sledování, kontrolu a řízení.

Rozvoj průmyslové výroby na konci 19. a počátku 20. století přináší potřebu zajištění kvality výroby „ve velkém“, tedy nikoli kvality jednotlivých výrobků při kusové výrobě, ale kvalitu velkého množství výrobků v sériové výrobě. K dosažení kvalitní výroby vedly dvě cesty: jednak zajištění kvality technologickými prostředky a výběrem materiálu, jednak kontrola kvality, která by vyloučila z produkce nevyhovující výrobky. Druhá cesta znamenala rozvoj takzvané „statistické přejímky“.

Rozvoj statistických teorií použitelných pro průmyslovou praxi přinesl růst výroby po první světové válce v minulém století. Americký profesor W. A. Shewhart ve třicátých letech položil základy kontroly výrobních procesů pomocí statistických metod ve své knize „Statistical Method from the Viewpoint of Quality Control“. Druhý výrazný skok v rozvoji statistických nástrojů pro kontrolu kvality nastal v období druhé světové války a v raně poválečném období především díky zbrojnímu průmyslu. Pozornost výrobců se soustředila na vlastní výrobu a technickou kontrolu vstupů a výstupů.

Od poloviny minulého století začaly výrazně narůstat požadavky zákazníků na výrobky a

jejich kvalitu. Bylo zřejmé, že výrobek, který pouze plní technologické parametry, nemusí být na trhu úspěšný a že zákazníci zohledňují i další kritéria jako hezký vzhled, spolehlivost, úspornost, komfort při užívání apod. Současně se stupňovaly požadavky na prodejní a poprodejní servis a návazné služby. Novou situaci a klima na světových trzích plně pochopili japonské strategové a manažeři. Ti jako první docenili teorii W. E. Deminga o komplexním přístupu k řízení kvality a s pomocí jeho metod prokázali, že toto chápání kvality je nejen konkurenční výhodou, ale i účinným nástrojem na cestě k prosperitě. S vývojem řízení kvality je spojena řada významných osobností, k nimž patří Walter A. Shewhart, William E. Deming, Joseph M. Juran, Armand V. Feigenbaum, Kaoru Ishikawa, Yoshi Tsurumi, Philip B. Crosby.

1.2 Historie DOE

- Klasický přístup - R. A. Fisher (20. léta 20. století). Návrh experimentů (DOE) je založen na statistických metodách, studiu společných efektů několika proměnných vlivů určením kombinace hodnot faktorů pro optimální výsledek
- Taguchiho přístup (40. léta 20. století) standardizoval a zjednodušil použití technik DOE, navrhl koncept zlepšování kvality ve všech fázích návrhu a výroby
- K významnému rozšíření metody DOE v USA a Evropě došlo až v 80. letech 20. století

1.3 Základní pojmy

Aby nedocházelo k nedorozuměním a nepřesným výkladům, zavedeme si několik základních pojmů, které se v souvislosti s navrhováním experimentů používají. Tyto pojmy definuje i norma ČSN ISO 3534-3 (viz [9]).

Proces je posloupnost operací či dějů, které systematicky přetvářejí vstupy na výstupy. Podle normy ČSN EN ISO 9000:2006 charakterizován jako „soubor vzájemně souvisejících nebo vzájemně působících činností, který přeměňuje vstupy na výstupy“. Výrobní proces tedy představuje posloupnost po sobě jdoucích činností, kde se snažíme za pomoci lidského zdroje, technických prostředků a technologie vytvořit produkt nebo službu, tím že přeměňujeme vstupní prvky na výstupní produkty. Výrobní proces můžeme také definovat jako činnost, při níž se uskutečňuje soubor pracovních, technologických a přírodních procesů, který mění tvar a jakost vstupního materiálu.

Experiment je test nebo série testů (pokusů), provedená za účelem zvýšení kvality produktu nebo procesu, případně zvýšení jejich efektivity. Slouží ke

- stanovení charakteristik procesu a jeho optimalizace
- vyhodnocení vlastností materiálů
- návrh a vývoj produktů
- stanovení tolerance komponent a vstupních veličin

Všechny experimenty jsou navržené experimenty, některé dobře, jiné špatně.

Návrh experimentu

- zkracuje dobu pro návrh a vývoj nových produktů
- zlepšuje fungování stávajících procesů
- zvyšuje spolehlivost a zlepšuje kvalitu výrobků
- zvyšuje robustnost výrobků a procesů
- umožňuje vyhodnocení různých variant, výběr komponent, nastavení parametrů a systémových tolerancí

Odezva – výstupní veličina, která je měřitelná, zpravidla spojitá.

Náhodný vliv – neznáme jeho příčiny, nelze jej odstranit. Způsobuje variabilitu, kterou lze měřit (experimentální chyba) a lze jej předvídat. Snaha je co nejvíce jej snížit.

Systematický vliv – je způsoben známými vlivy (vymežitelnými příčinami). Projevuje se například trendem, periodicitou, posunutím. Snažíme se jej popsat a kvantifikovat.

Faktory – vstupní veličiny, které mohou být kvalitativní (kategoriální) nebo kvantitativní (diskrétní, spojitě). Rozlišujeme hlavní, vedlejší a blokové.

Interakce – současné působení několika (alespoň dvou) faktorů.

Replikace – opakování zkoušek za (přibližně) stejných podmínek (úrovní faktorů). Umožňují měřit náhodnou variabilitu a oddělit ji od variability celkové.

Znáhodnění – stanovení pořadí zkoušek podle náhodného „zamíchání“. Do jisté míry může eliminovat vedlejší vlivy a zajišťuje vyšší míru „nezávislosti“ jednotlivých pokusů.

Uspořádání do bloků – slouží ke snižování náhodné variability (variability náhodné složky). V rámci bloku probíhají zkoušky za přibližně stejných experimentálních podmínek (ale při různých kombinacích úrovní faktorů). Blok často představuje jednu repliku experimentu.

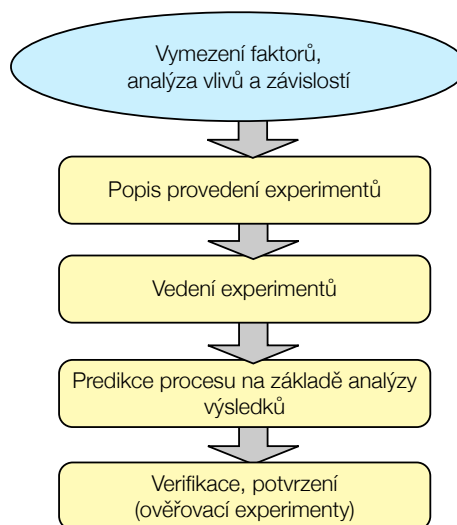
1.4 Návrh a analýza experimentu v pěti krocích

Hlavním cílem experimentu je odhalení závislostí a vztahů mezi různými veličinami, případně optimální (z určitého hlediska) nastavení vstupních parametrů procesu tak, aby jeho výstupy splňovaly požadovaná kritéria. Požadavek na provedení experimentu obvykle přichází z vlastní výroby, kde je potřeba vyřešit problémy spojené s nestabilitou, sníženou kvalitou výrobků, potřebou provést optimalizaci, nastavení parametrů a další. Naším cílem je obvykle nalézt příčiny neshod nebo zjistit optimální nastavení výrobních podmínek. Požadavky mohou přicházet také z oddělení marketingu či hodnocení kvality.

Návrh a analýzu experimentu obvykle provádíme v pěti na sebe navazujících krocích:

- Analýza procesu
- Návrh experimentu
- Provedení zkoušek
- Analýza výsledků
- Závěry

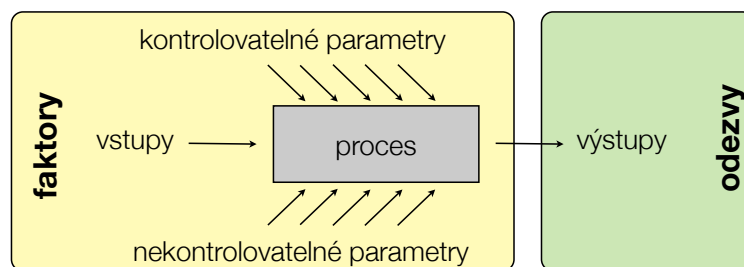
Tyto kroky by měly být provedeny vždy všechny a v daném pořadí. Často se tyto kroky cyklicky opakují a experiment se provádí víceúrovňově. V prvním kole provedeme hrubou analýzu a vyčleníme nejvlivnější faktory a později návrh zjemňujeme upřesňováním okruhu vlivných faktorů a hledáním jemnějších závislostí.



Obr. 1.1: Návrh a analýza experimentu v pěti krocích.

1.4.1 Analýza procesu

Návrh a analýza experimentů spočívá v pochopení vlivu různých proměnných na ostatní proměnné. Cílem je nalézt vztah mezi příčinami a následkem, v matematické terminologii je to vztah mezi nezávislými proměnnými a závislou proměnnou, která nás zajímá. Závisle proměnnou zde budeme nazývat *odezvou* a nezávisle proměnné se nazývají *faktory*. K určení nejvhodnější odezvy a identifikaci možných vlivných faktorů by měl být vytvořen tým složený z odborníků v různých oborech, týkajících se produktu nebo procesu. Týmový přístup podporuje synergii, která dává bohatší soubor podnětů ke studiu, a tedy i komplexnější experiment.



Obr. 1.2: Schematické znázornění výrobního procesu.

K identifikaci důležitých faktorů, které ovlivňují systém se používají takzvané *screeningové experimenty*. Tyto pokusy jsou prováděny na základě výzkumu a s využitím předchozí znalosti systému. Z velké skupiny potenciálních faktorů vyčleňují důležité faktory a zaměřují pozornost na klíčové faktory, které vyžadují další podrobnou analýzu.

1.4.2 Návrh experimentu

Průběh experimentu je důležité pečlivě naplánovat ještě před začátkem vlastního testování a sběrem dat. V této fázi je třeba mít zformulované důkladné a přesné určení cílů, jaká je třeba provést vyšetření, vyhodnocení potřebného času a zdrojů k dosažení cíle a integrovat předchozí znalosti k provedení vlastních experimentů. Pečlivě plánované experimenty vždy vedou k lepšímu pochopení produktu nebo procesu.

Experimenty probíhají při různých hodnotách faktorů, tzv. úrovních. Při každé provedené zkoušce budeme sledovat odezvu na nějakou, často předem stanovenou, kombinaci úrovní sledovaných faktorů. Tyto zvolené kombinace úrovní sledovaných faktorů se označují jako „ošetření“. Pokud je pro každé ošetření proveden stejný počet pozorování odezvové veličiny, označujeme návrh experimentu jako *homogenní*. Opakovaná pozorování při stejném ošetření jsou potom označována jako *replikace*.

Počet ošetření experimentu je stanoven na základě počtu úrovní sledovaného faktoru v experimentu. Například, pokud provádíme experiment, v němž sledujeme dva faktory, každý s konečným počtem úrovní. Bude-li mít první faktor r úrovní a druhý s úrovní, potom bude zřejmě potřeba provést pozorování při $r \times s$ kombinacích. Pokud jsou pozorovány všechny možné kombinace úrovní faktorů, budeme experiment označovat jako *úplný faktoriální*. Pokud pozorujeme pouze některé kombinace úrovní faktorů, řekneme že experiment je *dílčí faktoriální*. V případě úplných faktoriálních návrhů experimentů je třeba prověřit všechny faktory a jejich vzájemné interakce, zatímco u dílčích faktoriálních návrhů experimentů některé nebo všechny interakce chybí, protože ne všechny kombinace jsou pozorovány. Dílčí faktoriální návrhy se často používají jako screenin-

gové experimenty na počátku experimentování. V tomto stádiu máme na základě úvodní analýzy procesu vyčleněný velký počet možných vlivů na odezvu a musíme rozhodnout o významnosti jejich vlivu. Přitom očekáváme, že většina hlavních efektů a jejich interakcí bude nevýznamná a pro další experimenty je budeme moci vyloučit.

Obecně lze říci, že větší počet faktorů lépe popisuje model a můžeme zkoumat vliv více interakcí, na druhou stranu se návrh komplikuje velkým množstvím měření která vedou ke složité a nepřehledné analýze.

Norma ČSN ISO 3534/3 (Navrhování experimentů, 1993) rozlišuje tyto základní typy experimentů:

- Jednofaktorové experimenty
- Úplné vícefaktoriální experimenty (znáhodnění, uspořádání do bloků)
- Snížení počtu zkoušek (latinské, řeckolatské čtverce pro vyloučení vlivu rušivých faktorů)
- Dílčí faktoriální návrh (počáteční vyhledávání vlivných faktorů)
- Hierarchický návrh (vyhledání největších zdrojů variability)
- Optimální návrhy (optimální odezvové plochy)
- Taguchiho ortogonální návrhy (robustní návrhy)

Jednofaktorové experimenty

To jsou návrhy, kde je předmětem šetření pouze jediný faktor a cílem je určit, zda se odezva liší významně při různých úrovních tohoto faktoru. Faktor může být buď kvalitativní nebo kvantitativní. V případě kvalitativních faktorů (např. různí dodavatelé, různé materiály, apod.) není možná extrapolace a testování se provádí na jednotlivých úrovních. Vliv na odezvu může být určen pouze pro tyto úrovně. Na druhou stranu, je-li faktorem kvantitativní veličina (jako je teplota, napětí, zatížení, atd.), pro výpočet efektů je zpravidla k dispozici vždy dostatek údajů, neboť lze využít odhady a předpovědi. K vyhodnocení lze použít modely vícenásobné lineární regrese a další statistické metody.

Úplné vícefaktoriální návrhy

Ve úplném vícefaktoriálním experimentu je při každé zkoušce uvažován vliv několika faktorů současně. Stejně jako v jednofaktorovém experimentu, i zde rozlišujeme kvalitativní a kvantitativní faktory. Cílem těchto návrhů je identifikovat faktory, které mají významný vliv na odezvu,

stejně jako zkoumat vliv všech jejich interakcí (v závislosti na použité konstrukci experimentu). Každý každý faktor může mít různý počet úrovní.

Dvouúrovňové úplné faktoriální návrhy

Z hlediska volby počtu úrovní sledovaných faktorů je nejjednodušším (a také nejčastějším) případem dvouúrovňový experiment, kdy uvažované faktory sledujeme pouze ve dvou úrovních. Takovýmito experimenty jsme však schopni odhalit pouze lineární závislosti, ale už ne závislosti složitější. Máme-li pochybnosti o lineární závislosti odezvy a faktoru, volíme tři úrovně, vyjíměčně více, neboť potom se návrh stává velmi komplikovaný a vyžaduje velmi mnoho měření.

Dílčí faktoriálové experimenty

Zatímco v úplných experimentech je při větším počtu faktorů a jejich úrovní počet sledovaných kombinací zpravidla vysoký, tyto dílčí návrhy dovolují snížit počet kombinací úrovní faktorů na zlomek počtu, potřebného při úplném návrhu. V takových případech však nejsme schopni vyhodnotit efekty všech možných interakcí.

Taguchiho ortogonální návrhy

Taguchiho ortogonální experimenty jsou relativně jednoduché konstrukce. Používají se k odhadu hlavních efektů s použitím jen několika kombinací úrovní faktorů. Tyto návrhy jsou určeny jak pro dvouúrovňové faktoriální experimenty, tak i pro prozkoumání hlavních účinků faktorů s více než dvěma úrovněmi bez interakcí.

Taguchiho postup umožňuje nastavení těchto faktorů k dosažení požadovaného cíle. V závislosti na produktu nebo procesu, který je předmětem šetření tohoto cíle může být buď maximalizovat, minimalizovat nebo dosažení cílové hodnoty odezvy.

Optimální odezвовé plochy

Jedná se o speciální návrhy, které se používají k nastavení úrovní faktorů pro dosažení optimální hodnoty odezvy.

Spolehlivostní DOE

Jedná se o zvláštní kategorii, zahrnující klasické typy experimentů, jako jsou například dvouúrovňové experimenty, v kombinaci se spolehlivostními metodami. Cílem je zkoumat vliv různých faktorů na životnost přístroje. Ve spolehlivostním DOE je odezvou délka života zkoumaného objektu

(např. věk, mlé, cykly, atd.). Data mohou obsahovat cenzorovaná pozorování (závěsy, interval data).

1.4.3 Provedení zkoušek

Sebelepší návrh je bez pečlivého a důsledného provedení vlastních zkoušek a měření samoúčelný.

Pro jeho úspěšnou realizaci je potřeba věnovat velkou pozornost především

- dodržení zachování stejných podmínek okolního prostředí (měření v blocích),
- přesné nastavení kombinací úrovní sledovaných faktorů,
- respektovat navržené znáhodnění v experimentu, i za cenu jistých komplikací (například při opakovaném nastavování úrovní některého z faktorů),
- maximálnímu vyloučení šumových faktorů, jako je například vnější teplota, denní doba, lidský faktor a další.

Zvláště při velkém počtu měření je často nemožné provést všechna v jednom dni a za naprosto stejných podmínek. Pokud návrh experimentu s touto situací explicitně nepočítá (například rozdělením měření do bloků), je třeba věnovat velkou pozornost zachování alespoň přibližně stejných podmínek. Důležitá je randomizace. Ta má za cíl právě minimalizaci vlivu nesledovaných potenciálních faktorů, které navíc často nemůžeme ani ovlivnit. Randomizace neboli znáhodnění stanoví přesné pořadí měření a musí být součástí plánu experimentu.

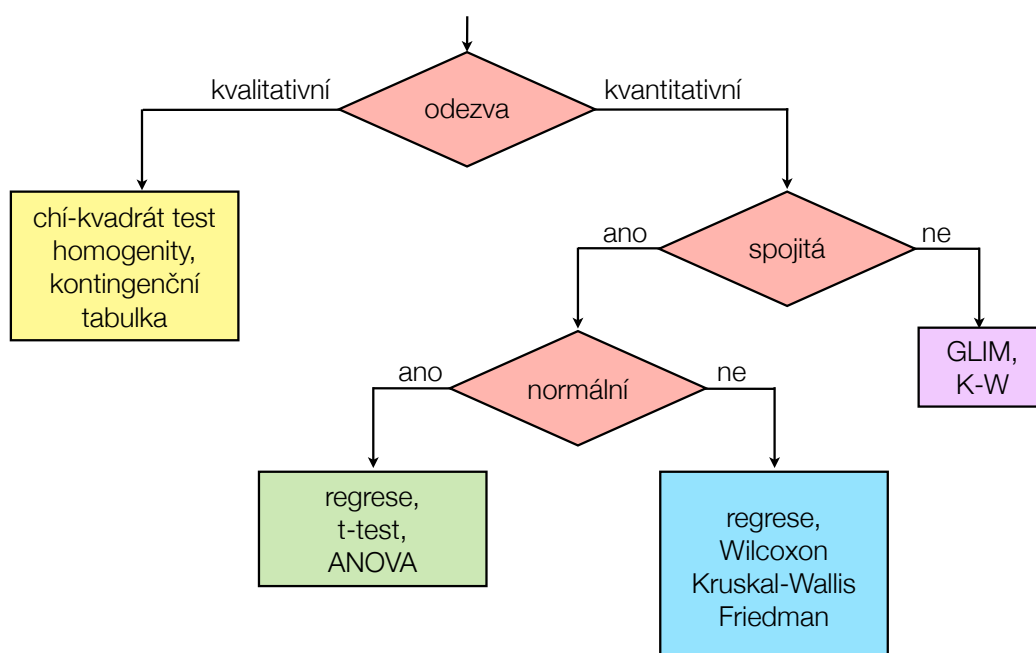
Samozřejmostí je dodržení základních metrologických zásad. Sem patří aplikace analýzy metrologického systému (MSA) před zahájením vlastních zkoušek. Tímto krokem zajistíme, že metrologické prostředky i užití metody měření jsou vhodné pro měření sledovaných faktorů a odezvy výrobního procesu.

1. Ověříme, zda veškeré metrologické prostředky jsou řádně vedeny v rámci metrologického řádu organizace. Tím máme na mysli, že metrologické prostředky podléhají řádné evidenci a periodické kontrole, v jejímž průběhu jsou ověřovány, respektive recalibrovány.
2. Dalším krokem je ověření systému měření. V průběhu tohoto kroku ověřujeme, zda při použití zvoleného měřidla a metody měření v konkrétním prostředí pro různé operátory získáváme naměřené hodnoty s minimální odchylkou od hodnoty konvenčně správné.

Metodika MSA ověřuje různé vlivy na variabilitu měření a v některých případech dokáže popsat příspěvky jednotlivých složek variability. Tyto složky obvykle dělíme na příspěvek operátora, měřidla, metody a prostředí.

1.4.4 Analýza výsledků

Vzhledem k tomu, že hodnoty odezvy při různých kombinacích úrovní sledovaných faktorů nikdy dopředu neznáme, považujeme odezvu vždy za náhodnou veličinu. Jednotlivá měření si můžeme představit jako jakýsi „náhodný výběr“ z možných realizací. K analýze výsledků měření potom používáme metody matematické statistiky. Tyto metody je třeba volit jednak v závislosti na charakteru odezvy, jednak na jejím pravděpodobnostním chování. Zde máme na mysli především její předpokládané rozdělení pravděpodobnosti. Zjednodušeně lze volbu metody pro vyhodnocení výsledků navržených experimentů zobrazit v rozhodovacím stromě (viz Obr. 1.3):



Obr. 1.3: Volba metody pro vyhodnocení výsledků měření v experimentu.

Stanovení metody pro analýzu výsledků měření je součástí návrhu experimentu. Pro návrháře experimentu a pro experimentátora je důležité znát tyto metody a především předpoklady pro jejich použití. Zároveň je důležité uvědomit si, co znamená porušení těchto předpokladů pro konečné závěry. Základní přehled některých matematicko-statistických metod je předmětem jiného kurzu a zde je uveden v příloze. Aplikaci těchto metod zpravidla provádíme prostřednictvím počítačového programu, což ovšem neznamená, že nemusíme těmto metodám rozumět a znát jejich principy. Naopak, mechanické používání softwarových prostředků bez znalosti v nich obsažených metod je velmi „nebezpečné“ a může být zavádějící. Navíc, tyto znalosti uplatníme při interpretaci výsledků.

1.4.5 Závěry

Tato poslední fáze zahrnuje potvrzení nejlepšího nastavení faktorů. Může představovat několik navazujících experimentů pro potvrzení, že systém funguje podle přání a všechny cíle byly splněny.

Kapitola 2

Stochastické repetitorium . . .

Jednou ze základních vlastností přírody a světa kolem nás je variabilita. Variabilita způsobuje, že provedeme-li dvakrát po sobě stejnou činnost, pokaždé bude její výsledek trochu jiný. Dokonce budeme-li pozorovat opakovaně tentýž jev, dostaneme odlišné výsledky. Odchytky se vyskytují v přírodě, ať už jde o pevnost v tahu určitého výrobku z oceli, obsah kofeinu v energetických nápojích, nebo vzdálenost, kterou urazí vaše vozidlo za den. Variabilitu lze také zaznamenat při měření různých průběhů procesu, probíhajících za stejných podmínek, při stejně nastavených parametrech a za stejného působení vnějších faktorů. Přírozené změny, které se vyskytují v procesu, přestože jsou všechny podmínky zachovány na stejné úrovni, často nazýváme *náhodným šumem*. Při provádění experimentů, při nichž zkoumáme vliv různých faktorů na proces, je důležité dokázat odlišit změny v procesu způsobené těmito faktory od změn způsobených náhodným šumem. K tomuto účelu lze použít řadu statistických metod. Proto je i v této učebnici věnována velká část právě náhodným veličinám, pravděpodobnostnímu rozložení a statistickým metodám testování hypotéz. Je důležité získat jasnou představu o pravděpodobnostním chování sledovaných veličin (faktorů), abychom mohli určit, zda jejich vliv na odezvu je významný či nikoliv. Proto klademe důraz na testování hypotéz které najde přímé použití při analýze a vyhodnocení navržených experimentů. Při vyhodnocování způsobu ovlivnění odezvy sledovanými faktory a kvantifikaci tohoto vlivu se uplatní metody regresní analýzy, která je také součástí tohoto repetitoria.

2.1 Náhodná veličina

2.1.1 Rozdělení náhodné veličiny

Pod pojmem **(reálná) náhodná veličina** budeme rozumět reálnou veličinu X , kterou jsme schopni pozorovat nebo měřit, ale jejíž hodnoty nejsme schopni určit před konkrétním měřením či pozorováním. V jistém smyslu si náhodnou veličinu můžeme představovat jako číselnou reprezentaci výsledků náhodného pokusu. Při aplikaci matematicko-statistických metod rozlišujeme mezi vlastní náhodnou veličinou, kterou zpravidla označujeme velkými písmeny a jejími realizacemi, které označujeme obvykle malými písmeny. Hodnoty náhodné veličiny X neznáme a pouze o nich víme, že leží v nějaké známé množině, zatímco její realizace x jsou konkrétní naměřené či napozorované hodnoty. Práce s náhodnými veličinami se tak výrazně odlišuje od práce s matematickými funkcemi jako je sinus, exponenciála nebo mocnina. Na rozdíl od matematických funkcí nelze například nakreslit graf náhodné veličiny. Nelze stanovit její průběh nebo limitu. Lze pouze stanovit její **pravděpodobnostní charakteristiky**.

Budeme-li náhodnou veličinu měřit či pozorovat opakovaně za přibližně stejných podmínek, může tato nabývat různých hodnot reálných čísel. Míru očekávání, že hodnota náhodné veličiny bude ležet v intervalu I , nazýváme *pravděpodobností* tohoto intervalu a budeme ji zapisovat formulí $P(X \in I)$. Pravděpodobnost splňuje několik základních pravidel, takzvaných *axiomů*:

- $P(X \in \emptyset) = 0$, kde \emptyset je prázdná množina,
- Jsou-li A a B dva disjunktní intervaly, potom $P(X \in A \cup B) = P(X \in A) + P(X \in B)$,
- $P(X \in A^C) = 1 - P(X \in A)$, kde A^C je doplněk intervalu A v množině reálných čísel.

Uvažujme dvě náhodné veličiny X a Y . Je-li $P(Y = y) > 0$ pro nějaké $y \in R$, potom výraz

$$P(X = x | Y = y) = \frac{P(X = x \wedge Y = y)}{P(Y = y)}$$

nazveme *podmíněnou pravděpodobností toho, že veličina X nabyde hodnoty x za podmínky, že veličina Y nabyla hodnoty y* . Pomocí podmíněné pravděpodobnosti definujeme tzv. *stochastickou nezávislost* dvou náhodných veličin: veličiny X a Y jsou stochasticky nezávislé právě když pro libovolné hodnoty $x, y \in R$ platí $P(X = x \wedge Y = y) = P(X = x) \cdot P(Y = y)$.

Stochastická nezávislost náhodných veličin je důležitou podmínkou řady matematicko-statistických metod a postupů. Její diagnostika však bývá často velmi složitá a obtížná.

Funkce $F(x)$ definovaná vztahem

$$F(x) = P(X \leq x)$$

se nazývá **distribuční funkce** náhodné veličiny X . (Tento zápis je zjednodušen vynecháním složené závorky u jevu $\{X \leq x\}$.) Základní vlastnosti této funkce jsou shrnuty v následujícím výčtu:

- (1) $0 \leq F(x) \leq 1$ pro každé $x \in R$,
- (2) $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$,
- (3) $\lim_{h \rightarrow 0+} F(x+h) = F(x)$, tj. F je zprava spojitá¹,
- (4) je-li $x_1 \leq x_2$, potom $F(x_1) \leq F(x_2)$, tj. F je neklesající.

Tyto vlastnosti jsou též postačující pro to, aby daná funkce $F(x)$ byla distribuční funkcí nějaké náhodné veličiny. Často se používá další vlastnost: pro reálná čísla $a \leq b$ platí

$$\begin{aligned} P(a < X \leq b) &= P(\{X \leq b\} \cap \{X > a\}) = P(\{X \leq b\} - \{X \leq a\}) \\ &= P(X \leq b) - P(X \leq a) = F(b) - F(a). \end{aligned}$$

Distribuční funkce $F(x)$ se nazývá **diskrétní**, existuje-li konečná nebo spočetná posloupnost bodů $\{x_i\}$ a posloupnost kladných čísel p_i splňujících $\sum_i p_i = 1$ takové, že

$$F(x) = \sum_{i: x_i \leq x} p_i$$

pro libovolné reálné číslo $x \in R$. Diskrétní distribuční funkce má schodovitý tvar, se skoky velikosti p_i v bodech x_i . Má-li náhodná veličina X diskrétní distribuční funkci, tj. $p_i = P(X = x_i)$, říkáme, že X má **diskrétní rozdělení pravděpodobnosti**, stručně **diskrétní rozdělení**.

Distribuční funkce $F(x)$ se nazývá **absolutně spojitá**, jestliže existuje spojitá nezáporná funkce $f(x)$ nazývaná **hustota pravděpodobnosti**, stručně **hustota**, taková, že

$$F(x) = \int_{-\infty}^x f(t) dt$$

¹V některých učebnicích se můžete setkat s poněkud jinou definicí distribuční funkce: $F(x) = P(X < x)$ (s ostrou nerovností). Takováto distribuční funkce potom bude spojitá zleva.

pro každé $x \in R$. Má-li náhodná veličina X absolutně spojitou distribuční funkci, říkáme, že X má **spojité rozdělení pravděpodobnosti**, stručně **spojité rozdělení**.

Hustota pravděpodobnosti $f(x)$ musí splňovat rovnost $\int_R f(x)dx = 1$. Existuje-li derivace F' distribuční funkce bodě x , potom je $F'(x) = f(x)$. Pro $a, b \in R, a < b$, platí

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t)dt$$

Tedy pravděpodobnost toho, že spojitá náhodná veličina bude mít hodnoty v nějakém intervalu $\langle a, b \rangle$ je tedy plocha pod křivkou hustoty nad intervalem $\langle a, b \rangle$.

Pravděpodobnostní vlastnosti náhodné veličiny jsou plně popsány její distribuční funkcí. Pomocí distribuční funkce lze určit rozdělení pravděpodobnosti $\{p_i\}_{i=1}^{\infty}$ v případě diskrétní náhodné veličiny nebo hustotu $f(x)$ v případě spojitě náhodné veličiny. To platí i obráceně: známe-li rozdělení pravděpodobnosti $\{p_i\}_{i=1}^{\infty}$ nebo hustotu $f(x)$, můžeme najít distribuční funkci $F(x)$.

2.1.2 Charakteristiky náhodných veličin

Kromě základních pravděpodobnostních charakteristik náhodné veličiny, kterými je její rozdělení pravděpodobnosti, distribuční funkce či hustota, se často používají i další, odvozené charakteristiky. Mezi ně patří především **momenty** a **kvantily**.

Z momentových charakteristik jsou nejznámější střední hodnota jako míra polohy a rozptyl (resp. směrodatná odchylka) jako míra rozptýlenosti, šikmost a špičatost. Z kvantilů jsou nejznámější minimum a maximum, dolní a horní decily, kvantily, medián. Z kvantilů jsou odvozeny i kritické hodnoty pro různé testy statistických hypotéz.

Střední hodnota $E(X)$ náhodné veličiny X s diskrétním resp. spojitým rozdělením je hodnota

$$E(X) = \sum_{i=1}^{\infty} x_i p_i, \quad \text{resp.} \quad E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Často budeme psát stručněji pouze EX (bez závorek). Střední hodnota je ve skutečnosti první obecný moment náhodné veličiny X a často bývá také označována jako *očekávaná hodnota*.

Rozptyl $Var(X)$ náhodné veličiny X (s diskrétním nebo spojitým rozdělením) je hodnota

$$Var(X) = E(X - E(X))^2.$$

Druhá odmocnina rozptylu je nazývána **směrodatnou odchylkou** náhodné veličiny X a bu-

deme ji obvykle označovat $\sigma(X)$. Rozptyl je ve skutečnosti tzv. *druhý centrální moment náhodné veličiny* X .

V případě, že neznáme rozdělení pravděpodobnosti náhodné veličiny, odhadujeme tyto charakteristiky pomocí naměřených hodnot při náhodném výběru². Máme-li k dispozici n měření náhodné veličiny X , tedy hodnoty x_1, x_2, \dots, x_n , potom střední hodnotu nejlépe odhadneme pomocí aritmetického průměru (tzv. *výběrového průměru*):

$$E(X) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Pro odhad rozptylu používáme zpravidla dvě varianty: *výběrový rozptyl* s^2 při malých rozsazích výběru (pro malá n) nebo odhad *rozptylu základního souboru* $\hat{\sigma}^2$ pro velká n :

$$s_X^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right), \quad \text{resp.} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2.$$

Vztah mezi střední hodnotou a rozptylem popisuje takzvaná „Čebyševova nerovnost“:

Čebyševova nerovnost: Pro náhodnou veličinu X s konečnou střední hodnotou μ a rozptylem σ^2 platí pro libovolné reálné číslo $\epsilon > 0$ nerovnost

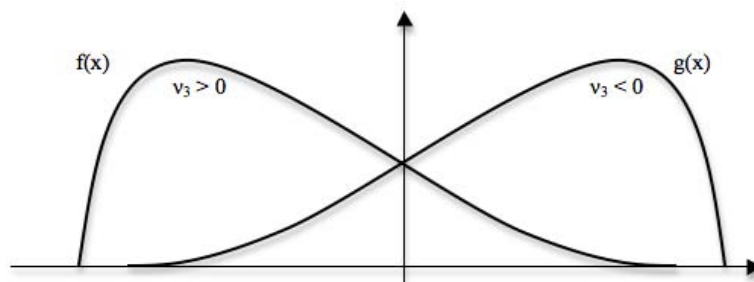
$$P(|X - E(X)| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}. \quad (2.1)$$

Často se pracuje s takzvanými **normovanými momenty**. Jsou to vlastně obecné momenty **normované náhodné veličiny** $U = \frac{X - E(X)}{\sigma(X)}$. Z těchto momentů jsou důležité ν_3 a ν_4 , které popisují tvar jejího rozdělení. Normovaný moment ν_3 se nazývá **koeficient šikmosti** a je mírou symetrie rozdělení. Koeficient šikmosti je roven nule například pro náhodné veličiny, jejichž rozdělení pravděpodobnosti je symetrické kolem střední hodnoty, je kladný pro jednovrcholové hustoty šikmé zprava (obr.2.1), naopak záporný pro jednovrcholové hustoty šikmé zleva.

Normovaný moment ν_4 je nazýván **koeficientem špičatosti** nebo také **kurtoze** a je mírou toho, jak rychle klesá pravděpodobnost extrémních hodnot (směrem k $-\infty$ nebo do $+\infty$).

Při studiu chování náhodné veličiny si zpravidla klademe otázku: *jaká je pravděpodobnost α , že sledovaná náhodná veličina X nepřekročí předem danou hodnotu x ?* Často je však kladena i opačná otázka: *jakou hodnotu x nepřekročí sledovaná náhodná veličina s předem danou pravděpodobností α ?* Odpověď nám dávají **kvantily rozdělení náhodné veličiny**, které tvoří

²Náhodný výběr znamená, že měříme sledovanou veličinu opakovaně, za přibližně stejných podmínek a nezávisle na předchozích naměřených hodnotách. Tím by měla být zajištěna tzv. *reprezentativnost*, tedy zastoupení jednotlivých hodnot ve výběru odpovídající neznámému rozdělení pravděpodobnosti.



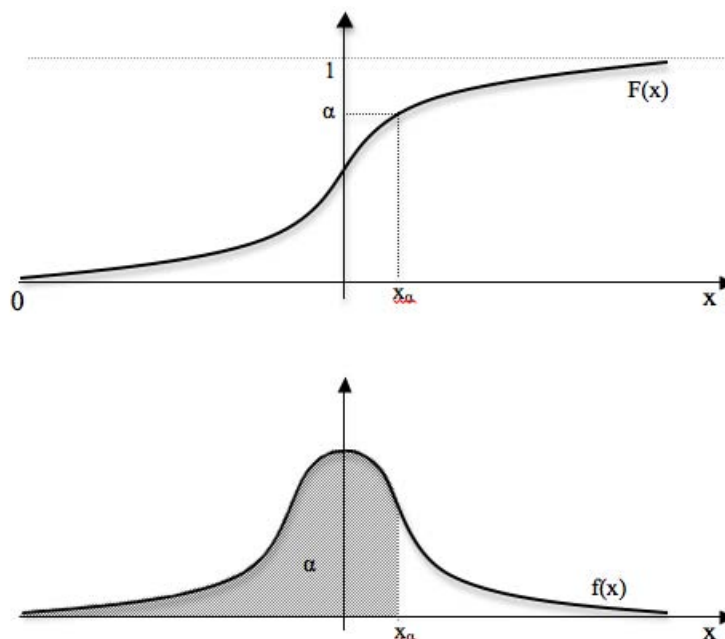
Obrázek 2.1: Hustoty nesymetrických rozdělání a jejich koeficienty šikmosti.

další důležitou skupinou charakteristik náhodné veličiny.

Mějme nějaké $0 < \alpha < 1$. Potom α -kvantilem náhodné veličiny X nazýváme takovou největší hodnotu x_α , pro kterou je

$$P(X \leq x_\alpha) \leq \alpha.$$

Má-li náhodná veličina X absolutně spojitou distribuční funkci $F(x)$, která je rostoucí pro ta x , pro která $0 < F(x) < 1$, potom existuje x_α takové, že $P(X \leq x_\alpha) = \alpha$ a je $x_\alpha = F^{-1}(\alpha)$, kde F^{-1} je inverzní funkce k F . Vyjádřeno pomocí hustoty f je $\int_{-\infty}^{x_\alpha} f(x)dx = \alpha$, viz obr.2.2.

Obrázek 2.2: Vztah mezi pravděpodobností α a α -kvantilem x_α .

Pravděpodobnost se často vyjadřuje v procentech. Potom budeme hovořit o $100\alpha\%$ -kvantilu náhodné veličiny. Podobně jako v případě momentů i zde platí tvrzení, že známe-li všechny α -kvantily náhodné veličiny X pro všechna $\alpha \in \langle 0, 1 \rangle$, potom máme úplnou informaci o jejím pravděpodobnostním chování.

Mezi všemi kvantily má významné postavení 50%-kvantil, který budeme označovat $x_{0,5}$ a budeme jej nazývat **medián** $Me(X)$ náhodné veličiny X . Medián se také někdy nazývá *prostřední hodnota* z hlediska pravděpodobnosti, neboť pravděpodobnost, že náhodná veličina X nabyde hodnoty menší než $Me(X)$ je rovna 0,5 což je stejná hodnota jako pravděpodobnost, že X bude mít hodnotu větší než $Me(X)$. Vedle střední hodnoty je to další takzvaná *charakteristika polohy* náhodné veličiny X .

Při analýze dat se často používají takzvané **kvartily**. To jsou 25%, 50% a 75% kvantily. Přitom 25%-kvantil se nazývá **dolní kvartil**, 50%-kvantil je již zmíněný medián a 75%-kvantil je **horní kvartil**. Spolu s minimem a maximem se těmito charakteristikám říká *pět Tukeyho charakteristik* podle zakladatele takzvané „průzkumové analýzy dat“, amerického statistika Johna Tukeye.

Ve statistice se dále pracuje s takzvaným *horním* a *dolním* 5%-kvantilem. To jsou 5%- a 95%-kvantily. Podobně se můžete setkat i s pojmem *horní* nebo *dolní decil* – tedy 10%- nebo 90%-kvantil.

2.1.3 Normální rozdělení pravděpodobnosti $N(\mu, \sigma^2)$.

Model normálního rozdělení má bohatou historii. Postupně byl objevován a opět zapomínám, až se trvale dostal do pořadí zájmu teorie pravděpodobnosti a především matematické statistiky. První, kdo popsal zvonovou křivku hustoty normálního rozdělení byl A. Moivre³ v roce 1733. K normálnímu modelu se dostal zobecněním binomického modelu při házení mincí. V té době mu nikdo nevěnoval zvláštní pozornost a křivka i rovnice upadly v zapomenutí. Až na přelomu 18. a 19. století ji znovu „objevili“ Gauss⁴ a Laplace⁵ při zkoumání astronomických měření. Byli postaveni před úlohu z mnoha měření, zatížených chybou, určit hodnotu, která se bude co nejvíce blížit skutečnosti. Odtud získal tento model přívlastek „model rozdělení chyb měření“ a odpovídající křivka hustoty se někdy též nazývá „Gaussova“. Třetí, kdo tento model objevil a zároveň první, kdo jej nazval „normálním“, byl Quételet⁶ v roce 1835. Normální křivku dostal

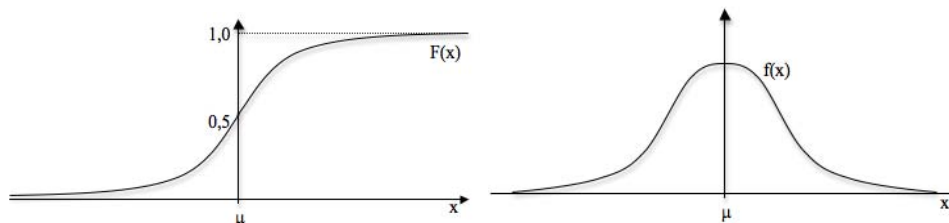
³Abraham de Moivre (1667–1754) byl francouzský matematik žijící větší část svého života v Anglii.

⁴Carl Friedrich Gauss (1777–1855) byl jeden z největších matematiků a fyziků všech dob. Zabýval se teorií čísel, matematickou analýzou, geometrií, geodézií, magnetismem, astronomií, optikou. Někdy bývá označován za „knížete matematiky“ nebo „největšího matematika od dob antiky“ – silně ovlivnil většinu oblastí svého oboru.

⁵Pierre Simon de Laplace (1749–1827) byl francouzský matematik, fyzik, astronom a politik; člen Francouzské akademie věd, královské společnosti v Londýně a Komise pro míry a váhy. Laplace je právem považován za jednoho z největších vědců vůbec. Zanechal monumentální dílo již svým rozsahem. Zabýval se matematickou analýzou, teorií pravděpodobnosti, nebeskou mechanikou, teorií potenciálu, zavedl pojem Laplaceovy transformace, užil tzv. Laplaceův operátor (v parciální diferenciální rovnici pro potenciál silového pole). Je autorem teorie o vzniku sluneční soustavy z rotující mlhoviny (Kantova-Laplaceova teorie) a mnoha dalších teorií a metod s mnoha aplikacemi

⁶Lambert Adolphe Jacques Quételet (1796–1874). Belgický vědec, jeden ze zakladatelů Královské statistické společnosti v Londýně

v souvislosti s měřením obvodu prsou 5738 skotských vojáků a představou jakéhosi „normálního“, neboli průměrného jedince. Od té doby si model normálního rozdělení začal budovat svoji pevnou pozici ve všech oblastech vědy.



Obrázek 2.3: Distribuční funkce a hustota normálního rozdělení.

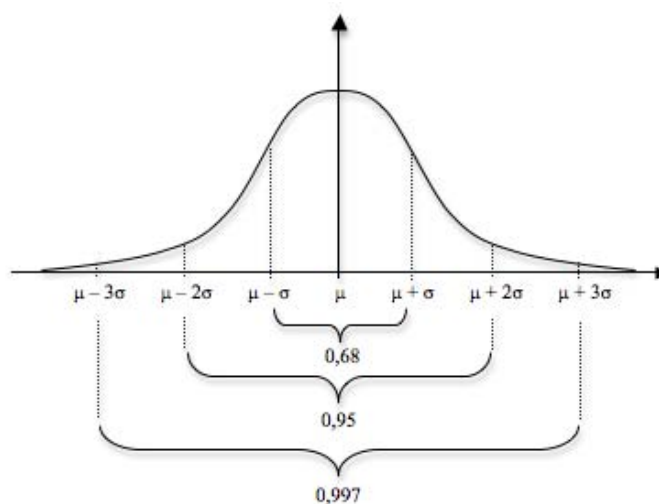
Hustota pravděpodobnosti: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $x \in R$

Distribuční funkce $F(x) = \int_{-\infty}^x f(t)dt$, $x \in R$

Základní charakteristiky: $E(X) = \mu$

$$Var(X) = \sigma^2$$

Poněkud nepříjemné v tomto modelu je to, že distribuční funkci, která je dána výše uvedeným integrálem, nelze vyjádřit konečnou analytickou formulí; její hodnoty se počítají pomocí distribuční funkce takzvaného *normovaného* normálního rozdělení (viz dále). V současné době to však není zásadní omezení, neboť řada programů (včetně tabulkového procesoru MS Excel) umějí distribuční funkci normálního rozdělení spočítat s dostatečnou přesností.



Obrázek 2.4: Vliv parametrů μ a σ normálního rozdělení na tvar křivky.

Parametry normálního rozdělení lze interpretovat jako *parametr polohy* $EY = \mu$ a *parametr měřítka* $VarY = \sigma^2$, vzhledem k symetrii rozdělení je μ též mediánem i modem. Význam směrodatné odchylky σ je ilustrován obr. 2.4, kde je znázorněna pravděpodobnost toho, že Y se

liší od střední hodnoty μ v absolutní hodnotě o méně než $k\sigma$, $k = 1, 2, 3$.

Často v praktických výpočtech používáme následující vlastnost normálního rozdělení: Má-li náhodná veličina X normální rozdělení pravděpodobnosti se střední hodnotou μ a rozptylem σ^2 ($X \approx N(\mu, \sigma^2)$), potom pro libovolné konstanty $a, b \in R, a > 0$ má veličina $Y = \frac{X-b}{a}$ opět normální rozdělení se střední hodnotou $\mu - b$ a rozptylem $(\frac{\sigma}{a})^2$, neboli patří ($Y \approx N(\mu - b, (\frac{\sigma}{a})^2)$).

Náhodná veličina Z má **normované normální rozdělení**, nebo též **standardní normální rozdělení**, ($Z \approx N(0, 1)$), je-li její hustota rovna $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, $z \in R$. Příslušná distribuční funkce se označuje obvykle symbolem Φ a lze ji vyjádřit jako

$$\Phi(z) = \frac{1}{2\pi} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt.$$

Hodnoty této funkce se počítají rozvojem v řadu a integrací člen po členu, nebo jsou uvedeny v takzvaných „Statistických tabulkách“.

n -tý obecný moment pro liché $n = (2k - 1)$ je vždy roven nule; tedy je $EZ^{(2k-1)} = 0, k \in N$. Pro n sudé, $n = 2k$ je $EZ^{2k} = \sqrt{\frac{2}{\pi}} \int_0^\infty z^{2k} e^{-\frac{z^2}{2}} dz$. Po substituci $\frac{z^2}{2} = t$ dostaneme⁷

$$EZ^{2k} = \sqrt{\frac{2}{\pi}} \int_0^\infty (2t)^{\frac{2k-1}{2}} e^{-t} dt = \frac{2^k}{\sqrt{(\pi)}} \Gamma(k + \frac{1}{2}) = (2k - 1)!! = 1.3.5 \dots (2k - 1)$$

Toto rozdělení se používá například tehdy, je-li třeba porovnat vlastnosti více náhodných veličin s různým normálním rozdělením. S takzvanou *normalizací* náhodné veličiny jsme se už setkali při definici *normovaných momentů* v odstavci 1.2. Jestliže má náhodná veličina X obecné normální rozdělení ($X \approx N(\mu, \sigma)$), vytvoříme normalizovanou náhodnou veličinu $Z = \frac{X-\mu}{\sigma}$. Tato veličina má normované normální rozdělení ($Z \approx N(0, 1)$). Vztah mezi hustotou $f(x)$ veličiny X a hustotou $\phi(z)$ veličiny Z je následující:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} = \frac{1}{\sigma} \phi(z).$$

neboli

$$f(x) = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right).$$

Distribuční funkci $F(x)$ veličiny X lze vyjádřit podobně pomocí $\Phi(z)$, neboť platí

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

⁷Symbol $(2k - 1)!!$ se používá k vyjádření „lichého“ faktoriálu, tedy součinu všech lichých čísel od 1 do $(2k - 1)$.

To lze snadno ověřit, dosadíme-li do integrálu pro $F(x)$ substituci $\frac{x-\mu}{\sigma} = z, dx = \sigma dz$.

Praktický význam normálního rozdělení vyplývá i z takzvané „centrální limitní věty“.

2.2 Testování statistických hypotéz

2.2.1 Test statistické hypotézy

Pod pojmem **statistické hypotézy** si budeme představovat jakékoli tvrzení o jevu statistické povahy. Například tvrzení o délce životnosti výrobku, o nezávislosti výsledku na použité metodě, tvrzení o pravděpodobnostním rozdělení sledované veličiny a podobně. Ověřování, zda hypotéza platí či nikoli, je předmětem statistického testování. Toto provádíme na základě pozorování (měření) nějakého výběru (experimentu).

Test statistické hypotézy H proti alternativní hypotéze A je rozhodovací pravidlo, podle něhož na základě realizace náhodného výběru rozhodujeme mezi dvěma tvrzeními - sledovanou hypotézou a doplňkovou, tzv. **alternativní hypotézou A** . Výsledkem našeho rozhodování je buď zamítnutí hypotézy H ve prospěch alternativy A či její nezamítnutí. Skutečnost, že hypotézu nezamítáme, neznamená, že naměřená data tuto hypotézu potvrzují, ale pouze to, že ji nevyvracejí. Toto rozhodovací pravidlo je určeno **testovou statistikou $T(X)$** a množinou ν , které říkáme **kritický obor**. Vlastní rozhodování potom probíhá pomocí indikátorové funkce

$$I_\nu(T(X)) = \begin{cases} 1 & \text{pokud } T(X) \in \nu, \\ 0 & \text{pokud } T(X) \notin \nu. \end{cases}$$

Pokud je hodnota indikátorové funkce rovna 1, tedy $T(X) \in \nu$ potom hypotézu H zamítáme. V opačném případě říkáme, že hypotézu nelze zamítnout. Naše úsilí přitom zaměříme na konstrukci testu, to znamená na určení testové statistiky T a kritického oboru ν , pomocí nichž budeme moci co nejlépe rozhodnout o zamítnutí hypotézy.

Příklad 2.2.1 *Při dodávce rezistorů je pro nás z hlediska použitelnosti rozhodující velikost odporu součástky. Výrobce udává nominální hodnotu, od níž se však většina naměřených hodnot liší. Jak rozhodnout, zda je pro nás dodávka přijatelná.*

Řešení: Naše měření však může podléhat náhodným vlivům. Kontrola dodávky spočívá ve stanovení rozhodovacího pravidla, kterým chceme otestovat hypotézu, že skutečný odpor je roven nominální hodnotě.

Při výše popsaném rozhodovacím pravidle se můžeme dopustit chyby dvěma způsoby. Buď budeme příliš přísní a zamítneme hypotézu, která platí - to je **chyba prvního druhu** - nebo naopak tuto hypotézu nezamítneme, i když je nesprávná - v tomto případě se jedná o **chybu druhého druhu**. Obě mohou mít nepříjemné důsledky, a proto budeme zřejmě za „lepší“ test považovat ten test, při kterém bude pravděpodobnost obou chyb co nejmenší. Přitom zpravidla čím menší bude pravděpodobnost chyby 1. druhu, tím větší bude pravděpodobnost chyby 2. druhu a naopak. V takovém případě nelze nalézt test minimalizující obě chyby současně. Proto postupujeme následujícím způsobem.

Při konstrukci testu požadujeme, aby pravděpodobnost chyby 1. druhu byl menší nebo rovna danému číslu α , kterému říkáme **hladina významnosti testu**. Přitom obvykle volíme $\alpha = 0,05; 0,01$ apod. Potom hledáme testovou statistiku $T(X)$ a kritický obor ν_α , tak aby

$$P(T(X) \in \nu_\alpha | H \text{ platí}) \leq \alpha$$

$P(T(X) \notin \nu_\alpha | H \text{ neplatí})$ byla minimální.

Kritický obor je zpravidla interval, ohraničený tzv. **kritickými hodnotami**. Test potom probíhá tak, že spočteme hodnotu testové statistiky, porovnáme ji s kritickými hodnotami, odpovídajícími hladině významnosti α , a rozhodneme o zamítnutí či nezamítnutí hypotézy.

V některých případech - především při testování pomocí počítače - se používá jiný postup. Spočte se hodnota testové statistiky a k ní nejmenší kritický obor, při kterém bychom ještě mohli na základě této hodnoty zamítnout hypotézu proti dané alternativě. Hladina významnosti, odpovídající tomuto kritickému oboru, se nazývá **p -hodnota**. Kdybychom volili hladinu významnosti větší, než je tato hodnota, mohli bychom ještě hypotézu zamítnout. Je-li tato p -hodnota příliš malá, hypotézu zamítáme. Například, spočteme-li pro daná data p -hodnotu rovnou 0,005 znamená to, že pro jakékoliv α větší než 0,005 bychom měli hypotézu zamítnout. Konkrétně pro $\alpha = 0,01$ už hypotézu zamítáme.

Širokou třídu testů tvoří testy hypotéz o parametrech pravděpodobnostního rozdělení. V tomto případě předpokládáme, že pravděpodobnostní rozdělení je určitého typu a závisí na neznámých parametrech. Předpokládejme, že neznámý parametr θ může nabývat hodnot z nějaké množiny Θ . Uvažujme hypotézu $H : \theta = \theta_0$. Alternativní hypotéza může být buď $A : \theta \neq \theta_0$, takzvaná **oboustranná alternativa** při této alternativě mluvíme o **oboustranném testu**), nebo $A_1 : \theta \leq \theta_0$ či $A_2 : \theta \geq \theta_0$ tzv. **jednostranné alternativy** (jimž odpovídají tzv. **jednostranné testy**).

Funkci $P_\nu(\theta)$, která každé hodnotě parametru $\theta \in \Theta$ přiřadí pravděpodobnost $P(T(X) \in \nu|\theta)$ nazýváme **silofunkcí testu**. Je to vlastně pravděpodobnost zamítnutí hypotézy H , má-li parametr hodnotu θ . Hodnotu silofunkce v bodě $\theta = \theta_1$ nazýváme **silou testu vzhledem k alternativě $\theta = \theta_1$** a používáme ji k pro hodnocení kvality testu.

2.2.2 Testy o parametrech normálního rozdělení

V první části tohoto odstavce uvedeme několik parametrických testů, které se používají při výběru $X = X_1, \dots, X_n$ z rozdělení $N(\mu, \sigma^2)$. Tyto testy se zabývají hypotézami o parametrech μ a σ . Přitom rozlišujeme několik případů:

Předpokládejme, že hodnotu σ známe. Budeme testovat nulovou hypotézu $H : \mu = \mu_0$ proti alternativě $A : \mu \neq \mu_0$ na hladině významnosti α . Test založíme na statistice $\frac{\bar{X} - \mu}{\sigma} \sqrt{n}$, která má, za předpokladu platnosti hypotézy, rozdělení $N(0, 1)$. Kritický obor je potom určen nerovností

$$\frac{|\bar{X} - \mu|}{\sigma} \sqrt{n} > u(1 - \frac{\alpha}{2})$$

kde $u(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$ -kvantil rozdělení $N(0, 1)$. Pokud bude hodnota výběrového průměru $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ získaná z pozorování x_1, \dots, x_n ležet v intervalu

$$\mu_0 - \frac{\sigma}{\sqrt{n}} u(1 - \frac{\alpha}{2}) < \bar{x} < \mu_0 + \frac{\sigma}{\sqrt{n}} u(1 - \frac{\alpha}{2})$$

hypotézu nezamítneme na hladině významnosti α . V opačném případě hypotézu zamítáme. *Poznámka:* Srovnajte tento výsledek s intervalem spolehlivosti v VI.3.2. Hypotézu H nezamítneme tehdy, když hypotetická hodnota μ_0 bude ležet v $100(1 - \alpha)\%$ intervalu spolehlivosti, zkonstruovaném na základě pozorování x_1, \dots, x_n .

Při jednostranné alternativě $A_1 : \mu < \mu_0$ resp. $A_2 : \mu > \mu_0$ bude kritický obor určen nerovností

$$\frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} \leq u(\alpha), \text{ resp. } \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} \geq u(1 - \alpha)$$

Jednovýběrový t -test. Nejčastějším případem je test hypotézy $H : \mu = \mu_0$ proti alternativě $A : \mu \neq \mu_0$ na hladině významnosti α při neznámé hodnotě σ . Neznámou hodnotou σ nahrazujeme jejím odhadem S a test založíme na statistice $\frac{\bar{x} - \mu_0}{S} \sqrt{n}$. Platí-li hypotéza H , má tato statistika Studentovo t -rozdělení o $(n - 1)$ stupních volnosti (viz V.4.6). Kritický obor je potom určen nerovností

$$\frac{|\bar{x} - \mu_0|}{S} \sqrt{n} > t_{n-1}(1 - \frac{\alpha}{2})$$

kde $t_{n-1}(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$ -kvantil rozdělení t o $n - 1$ stupních volnosti.

Pro jednostranné testy proti alternativám $\mu < \mu_0$ resp. $\mu > \mu_0$ bude kritický obor určen nerovnostmi

$$\frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sqrt{n} \leq -t_{n-1}(1 - \alpha) \quad \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sqrt{n} \geq t_{n-1}(1 - \alpha)$$

Pro výběry o velkém rozsahu n lze t -test použít (přibližně) i bez předpokladu normality výběru. Podle centrální limitní věty IV.2.4 má totiž výběrový průměr \bar{X} v limitě pro $n \rightarrow \infty$ normální rozdělení $N(\mu, \sigma^2)$ a tedy statistika $\frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sqrt{n}$ má přibližně Studentovo t -rozdělení o $(n - 1)$ stupních volnosti.

Chceme-li testovat hypotézu o rozptylu $H : \sigma^2 = \sigma_0^2$ proti alternativě $A : \sigma^2 \neq \sigma_0^2$ na hladině významnosti α , můžeme použít statistiku $\frac{(n-1)S^2}{\sigma_0^2}$, která má za platnosti hypotézy (viz V.4.6) rozdělení χ o $n - 1$ stupních volnosti. Označme s^2 hodnotu výběrového rozptylu $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$, získanou z pozorování x_1, \dots, x_n . V tomto případě je kritický obor pro oboustrannou alternativu $A : \sigma^2 \neq \sigma_0^2$ určen nerovnostmi

$$s^2 < \frac{\chi(n-1)^2(\frac{\alpha}{2})\sigma_0^2}{n-1} \quad \text{a} \quad \frac{\chi(n-1)^2(\frac{1-\alpha}{2})\sigma_0^2}{n-1} < s^2$$

zatímco pro jednostranné alternativy $A_1 : \sigma^2 < \sigma_0^2$ resp. $A_2 : \sigma^2 > \sigma_0^2$ dostaneme kritické obory

$$s^2 < \frac{\chi(n-1)^2(\alpha)\sigma_0^2}{n-1} \quad \text{resp.} \quad \frac{\chi(n-1)^2(1-\alpha)\sigma_0^2}{n-1} < s^2,$$

symbol $\chi(n-1)^2(\alpha)$ zde označuje α -kvantil chí-kvadrát rozdělení o $(n - 1)$ stupních volnosti.

Párový t-test. Sledujeme-li na jednom objektu dva podobné znaky zároveň, používáme náhodný výběr dvojic náhodných veličin $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$. O veličinách X_i a Y_i předpokládáme, že jsou **párově závislé**. Například měření vlastnosti materiálu před tepelným zpracováním a po něm na vybraných n vzorcích.

Předpokládejme, že $\{X_1, \dots, X_n\}$ je náhodný výběr z normálního rozdělení $N(\mu_X, \sigma^2)$ a Y_1, \dots, Y_n je náhodný výběr z rozdělení $N(\mu_Y, \sigma_Y^2)$. Veličiny X_i a Y_i mohou být párově závislé. Budeme testovat hypotézu o rovnosti středních hodnot $H : \mu_X = \mu_Y$ proti alternativě $A : \mu_X \neq \mu_Y$ na hladině významnosti α . V tomto případě budeme místo původně sledovaných veličin $(X_1, Y_1), \dots, (X_n, Y_n)$ pracovat s veličinami Z_1, \dots, Z_n , kde $Z_i = Y_i - X_i, i = 1, \dots, n$. Protože X_i a Y_i mají normální rozdělení, bude se i veličina Z řídit normálním rozdělením se střední hodnotou $\mu_Z = \mu_Y - \mu_X$ a rozptylem σ_Z^2 , o jehož vztahu k rozptylům σ_X a σ_Y nelze vzhledem k možné závislosti nic předpokládat. Rovnost středních hodnot X a Y je ekvivalentní nulovosti

střední hodnoty rozdílu Z . Pro aritmetický průměr platí $\bar{z} = \bar{x} - \bar{y}$ a hodnotu výběrového rozptylu s_Z^2 spočteme podle vztahu

$$s_Z^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - y_i - \bar{x} + \bar{y})^2.$$

K testu hypotézy $H : \mu_Z = 0$ na hladině významnosti α použijeme jednovýběrový t-test (viz VII.2.2), tedy při oboustranné alternativě $A : \mu \neq 0$ hypotézu H zamítneme, pokud

$$\bar{z} = -\frac{s_Z}{\sqrt{n}} t_{n-1} \left(1 - \frac{\alpha}{2}\right) \text{ nebo } \frac{s_Z}{\sqrt{n}} t_{n-1} \left(1 - \frac{\alpha}{2}\right) < \bar{z}.$$

Jsou-li veličiny X a Y nezávislé, používáme pro srovnání středních hodnot dvou výběrů **dvouvýběrový t-test**. Nechť X_1, \dots, X_n je výběr z rozdělení $N(\mu_X, \sigma_X^2)$ a Y_1, \dots, Y_m je výběr z rozdělení $N(\mu_Y, \sigma_Y^2)$ a tyto výběry jsou na sobě nezávislé. Rozlišujeme dva případy:

(1) Oba výběry mají **stejné rozptyly** $\sigma_X^2 = \sigma_Y^2$. Potom statistika

$$T = \frac{\bar{X} - \bar{Y}}{S} \sqrt{\frac{mn}{m+n}},$$

kde

$$S^2 = \frac{1}{m+n-2} [\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2]$$

má za platnosti nulové hypotézy $H : \mu_X = \mu_Y$ Studentovo t-rozdělení o $m+n-2$ stupních volnosti. Test uvedené hypotézy proti oboustranné alternativě $A : \mu_X \neq \mu_Y$ na hladině významnosti α lze tedy založit na nerovnosti

$$\frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} \geq t_{m+n-2} \left(1 - \frac{1}{2}\alpha\right).$$

Test hypotézy H proti jednostranným alternativám $A_1 : \mu_X \leq \mu_Y$, resp. $A_2 : \mu_X \geq \mu_Y$ na hladině významnosti α je založen na nerovnostech

$$\begin{aligned} \frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} &\geq t_{m+n-2}(1 - \alpha), \\ \frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} &\geq t_{m+n-2}(\alpha) = -t_{m+n-2}(\alpha)(1 - \alpha) \end{aligned}$$

(2) Oba výběry mají **různé rozptyly** $\sigma_X^2 \neq \sigma_Y^2$. Potom použijeme přibližný test, založený na statistice

$$\tilde{T} = \frac{\bar{X} - \bar{Y}}{\tilde{S}}, \text{ kde } \tilde{S}^2 = \frac{1}{n}s_X^2 + \frac{1}{m}s_Y^2.$$

Test hypotézy $H : \mu_X = \mu_Y$ proti oboustranné alternativě na hladině významnosti α lze založit na nerovnosti

$$\frac{|\bar{y} - \bar{x}|}{\tilde{S}} \geq \frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \frac{\alpha}{2}) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \frac{\alpha}{2}) \right]$$

na nerovnosti

$$\frac{\bar{y} - \bar{x}}{\tilde{S}} \geq \frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \alpha) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \alpha) \right]$$

při jednostranné alternativě $A_1 : \mu_X \leq \mu_Y$ a na

$$\frac{\bar{y} - \bar{x}}{\tilde{S}} \leq -\frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \alpha) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \alpha) \right]$$

při alternativě $A_2 : \mu_X \geq \mu_Y$.

Příklad 2.2.2 Při zpracování je třeba materiál zahřát na vysokou teplotu. Před zpracováním bylo vybráno náhodně 10 vzorků a změřena jejich tvrdost. Po zpracování bylo opět vybráno náhodně jiných 10 vzorků, na nichž byla změřena tvrdost. Naměřené hodnoty jsou v následující tabulce:

před	3,15	2,98	3,00	2,75	3,21	3,33	2,95	2,81	3,26	2,88
po	3,21	2,99	3,11	2,91	3,22	3,28	3,09	3,00	3,28	2,99

Testujte hypotézu, že se tvrdost materiálu vlivem zpracování nemění.

Řešení: Je $m = n = 10$. Spočteme $\bar{x} = 3,032$, $\bar{y} = 3,108$, $s_x^2 = 0,03875$, $s_y^2 = 0,018$. Za předpokladu, že rozptyl před i po zpracování zůstává stejný (naměřený rozdíl je nevýznamný), použijeme postup, popsáný v VII.2.6.a). Dostaneme hodnotu $s^2 = 0,02838$ a testové statistiky $T = 1,009$. Při hladině významnosti $\alpha = 0,05$ je $t_{18}(0,975) = 2,101$ a tedy hypotézu nelze zamítnout.

Příklad 2.2.3 Uvažujme stejnou úlohu jako v předchozím příkladu, pouze s tím rozdílem, že sledovaná veličina je měřena před i po zpracování na 10 vzorcích, které byly náhodně vybrány před začátkem experimentu. Naměřená data zůstávají stejná.

Řešení: V tomto případě je třeba vzít do úvahy závislost, která zde může být způsobena dalšími vlastnostmi vzorků. Proto použijeme párový t-test. Dostáváme $\bar{z} = \bar{x} - \bar{y} = -0,076$, $s_z = 0,07777$. Testová statistika zde bude mít hodnotu $T = 3,09$, kterou budeme srovnávat s číslem $t_9 = (0,975) = 2,262$. V tomto případě hypotézu zamítneme. Uvedené příklady ukazují, jaký vliv na

výsledek může mít tzv. *návrh experimentu*. Druhý případ lépe vystihuje skutečnost, že naměřená data nejsou nezávislá a bere do úvahy další možné vlivy, plynoucí z individuality vzorků.

Uvažujme dva nezávislé výběry: X_1, \dots, X_n z rozdělení $N(\mu_X, \sigma_X^2)$ a Y_1, \dots, Y_n z rozdělení $N(\mu_Y, \sigma_Y^2)$ můžeme provést tzv. **test shody rozptylů** neboli **F-test**. K testu hypotézy $H : \sigma_X^2 = \sigma_Y^2$ lze použít například statistiku

$$F = \frac{S_X^2}{S_Y^2}$$

Rozdělení statistiky F je podle V.4.6 a V.4.9 za předpokladu H rozdělením F o $(n-1)$ a $(m-1)$ stupních volnosti. Kritický obor pro test hypotézy H proti oboustranné alternativě $A : \sigma_X^2 \neq \sigma_Y^2$ na hladině významnosti α je určen nerovnostmi

$$\frac{s_X^2}{s_Y^2} \geq F_{n-1, m-1}(1 - \frac{\alpha}{2}) \text{ a } \frac{s_X^2}{s_Y^2} \leq F_{n-1, m-1}(\frac{\alpha}{2}) = \frac{1}{F_{n-1, m-1}(1 - \frac{\alpha}{2})}$$

kde $F_{n,m}(\alpha)$ je α -kvantil rozdělení $F(n, m)$. Tedy hypotézu zamítáme pro malá F blízká nule a pro velká F (při platnosti hypotézy by mělo být F blízké 1). Kritické obory pro test hypotézy H proti alternativám $A_1 : \sigma_X^2 \geq \sigma_Y^2$ $A_2 : \sigma_X^2 \leq \sigma_Y^2$ na hladině významnosti α jsou určeny nerovnostmi

$$\frac{s_X^2}{s_Y^2} \geq F_{n-1, m-1}(1 - \alpha) \text{ resp. } \frac{s_X^2}{s_Y^2} \leq F_{n-1, m-1}(\alpha) = \frac{1}{F_{n-1, m-1}(1 - \alpha)}$$

2.2.3 Testy dobré shody

Při vyhodnocení výsledků experimentu velmi často předpokládáme, že naměřené hodnoty jsou realizacemi náhodných veličin s normálním rozdělením. K prověření hypotézy o typu rozdělení, z něhož byl výběr pořízen používáme takzvané **testy dobré shody**. Jedná se tedy o hypotézy o *shodě* teoretického a empirického rozdělení.

Test hypotézy, že náhodný výběr pochází z rozdělení se spojitou známou distribuční funkcí $F_0(x)$, můžeme provést pomocí takzvaného **Kolmogorov-Smirnovova testu**. Tento test je založen na statistice

$$D = \sup_{x \in R} \left\{ |F_n(x) - F_0(x)| \right\},$$

kde empirická distribuční funkce $F_n(x)$ je definována v paragrafu V.3.6. Pomocí této definice

lze statistiku D zapsat také ve tvaru

$$D = \max_{1 \leq i \leq n} \left\{ \max \left[\left| F_0(X_{(i)}) - \frac{i}{n} \right|, \left| F_0(X_{(i)}) - \frac{i-1}{n} \right| \right] \right\}$$

Hypotéza $H : F(x) = F_0(x)$ bude zamítnuta na hladině významnosti α ve prospěch alternativy $A : F(x) \neq F_0(x)$ alespoň pro jedno x , jestliže $D \leq D_n(1 - \alpha)$, přičemž hodnoty $D_n(1 - \alpha)$ jsou malá n tabelovány viz [Ja] a pro velká n ($n > 100$) lze použít aproximaci (viz [Zv])

$$D_n(1 - \alpha) \cong \sqrt{-\frac{1}{2n} \ln \frac{\alpha}{2}}.$$

Tento test lze správně použít pouze pro takové hypotézy, které určují funkci F_0 , jednoznačně, včetně jejich parametrů.

χ^2 -test dobré shody vychází z třídniho rozdělení náhodného výběru. Nejprve tedy provedeme rozklad naměřených hodnot do disjunktních třídních intervalů pomocí zvoleného dělení $a_0 < a_1 < a_2 < \dots < a_k$ a spočteme četnosti n_i (viz V.2.1). Dále spočteme hypotetické pravděpodobnosti $p_i = F_0(a_i) - F_0(a_{i-1})$. Při volbě třídních intervalů se doporučuje dodržet zásadu aby teoretické četnosti np_i pro všechna i byly větší nebo alespoň rovny číslu 5. Hypotézu, že výběr je z rozdělení s distribuční funkcí $F_0(x)$ potom testujeme pomocí statistiky

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}$$

Při tomto postupu je třeba rozlišit dva případy:

a) Hypotézu určuje distribuční funkci jednoznačně, včetně parametrů. Potom má statistika χ^2 asymptoticky, to znamená přibližně pro velká n , rozdělení $\chi^2(k - 1)$. Bude-li tedy

$$\chi^2 \geq \chi_{k-1}^2(1 - \alpha)$$

kde $\chi_{k-1}^2(1 - \alpha)$ je $(1 - \alpha)$ -kvantil rozdělení $\chi^2(k - 1)$ zamítneme nulovou hypotézu $H : X_1, \dots, X_n$ pochází z rozdělení s distribuční funkcí $F_0(x)$ proti alternativě $A : \text{toto rozdělení je jiné}$, na hladině významnosti α

b) Teoretické četnosti p_i závisí na l neznámých parametrech. V takovém případě použijeme při výpočtu p_i odhady těchto parametrů. Přitom se sníží počet stupňů volnosti rozdělení statistiky χ^2 právě o počet odhadnutých parametrů na $(k - 1 - l)$. Kritický obor při hladině významnosti α potom bude dán nerovností

$$\chi^2 \geq \chi_{k-1-l}^2(1 - \alpha)$$

K ověření předpokladu normálního rozdělení lze použít **testy normality** založené na výběrových koeficientech **šikmosti** A_3 a **špičatosti** A_4 . Pro tyto statistiky (viz. V.3.4) platí následující vztahy:

$$E(A_3) = 0, \quad Var(A_3) = \frac{6(n-2)}{(n+1)(n+3)}$$

$$E(A_4) = -\frac{6}{n+1}, \quad Var(A_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

přičemž $\sqrt{n}A_3$ a $\sqrt{n}A_4$ mají při $n \rightarrow \infty$ přibližně normálního rozdělení. Test založený na šikmosti zamítne hypotézu o normalitě, pokud

$$\frac{|A_3|}{\sqrt{Var(A_3)}} \geq u\left(1 - \frac{\alpha}{2}\right)$$

test založený na špičatosti zamítne hypotézu o normalitě, pokud

$$\frac{|A_4 - E(A_4)|}{\sqrt{Var(A_4)}} \geq u\left(1 - \frac{\alpha}{2}\right)$$

kde $u(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$ -kvantil rozdělení $N(0, 1)$. Hladina významnosti obou testů je asymptoticky rovna α . Při ověřování normality je vhodné provést oba testy. Hypotézu nezamítneme teprve tehdy, pokud ji nelze zamítnout oběma testy zároveň. Tyto testy jsou v některých případech citlivější na porušení normality než χ^2 test, podrobnější informace nalezne čtenář v [An].

2.2.4 Některé neparametrické testy

Znaménkový test je test o hodnotě mediánu (viz V.3.7). Předpokládáme, že výběr X_1, \dots, X_n je z rozdělení se spojitou distribuční funkcí $F(x)$. Budeme testovat hypotézu $H : \widetilde{x}_{50} = x_0$ proti alternativě $\widetilde{x}_{50} \neq x_0$ na hladině významnosti α . Vytvoříme posloupnost rozdílů $X_1 - x_0, \dots, X_n - x_0$. Označme Z počet členů této posloupnosti s kladným znaménkem a m počet nenulových rozdílů. Z je náhodná veličina s binomickým rozdělením (viz II.3.2) s parametry m a $\frac{1}{2}$. Pro malá m lze tedy stanovit kritický obor pro dané α určením celého čísla c tak, aby byly splněny nerovnosti

$$\sum_{i=1}^c \binom{m}{i} \left(\frac{1}{2}\right)^m \leq \frac{\alpha}{2} < \sum_{i=1}^{c+1} \binom{m}{i} \left(\frac{1}{2}\right)^m$$

Hypotézu H potom zamítneme, pokud $Z < c$ nebo $m - c < Z$. Pro větší m lze využít aproximace binomického rozdělení rozdělením normálním (viz věta v IV.2.1) a kritický obor vyjádřit pomocí $(1 - \frac{\alpha}{2})$ -kvantilu rozdělení $N(0, 1)$ nerovností $\frac{|2Z - m|}{\sqrt{m}} \geq u(1 - \frac{\alpha}{2})$. Pro jednostranné alternativy vytvoříme kritický obor analogicky jako v VII.2.1.

Jednovýběrový Wilcoxonův test. Předpokládáme, že $\{X_1, \dots, X_n\}$ je výběr z rozdělení se spojitou distribuční funkcí $F(x)$, která je symetrická kolem mediánu \widetilde{x}_{50} (neboli $F(\widetilde{x}_{50} - x) =$

$1 - F(\widetilde{x}_{50} + x)$. Budeme opět testovat hypotézu $H : (\widetilde{x}_{50} = x_0)$ proti alternativě $(\widetilde{x}_{50} \neq x_0)$ na hladině významnosti α . Podobně jako v VII.4.1 i v tomto případě vytvoříme posloupnost rozdílů $X_1 - x_0, \dots, X_n - x_0$ a dále budeme počítat pouze s nenulovými rozdíly, jejichž počet označíme m . Tuto posloupnost uspořádáme vzestupně podle absolutních hodnot a označíme R_i^+ pořadí náhodné veličiny $|X_i - x_0|$. Sečteme-li pořadí R_i^+ pro všechny členy, pro které je $X_i - x_0 > 0$ a tento součet označíme S^+ , dostaneme statistiku, pro kterou za platnosti hypotézy platí

$$E(S^+) = \frac{m(m+1)}{4}, \quad Var(S^+) = \frac{m(m+1)(2m+1)}{24}$$

a pro velká m je její rozdělení přibližně normální. Proto budeme pracovat raději s normovanou veličinou $V = \frac{S^+ - E(S^+)}{\sqrt{Var(S^+)}}$. Hypotézu tedy zamítneme, pokud $|V| \geq v(1 - \frac{\alpha}{2})$. Pro malé hodnoty m jsou kritické hodnoty $v(1 - \frac{\alpha}{2})$ tabelovány, pro velká m lze použít kvantily rozdělení $N(0, 1)$.

Dvouvýběrový Wilcoxonův test slouží k testování hypotézy o shodě distribučních funkcí dvou výběrů. Necht' $\{X_1, \dots, X_n\}$ a $\{Y_1, \dots, Y_m\}$ jsou dva nezávislé výběry ze dvou spojitých rozdělení. Za platnosti hypotézy jsou tato rozdělení totožná a spojený výběr $\{X_1, \dots, X_n, Y_1, \dots, Y_m\}$ lze považovat za výběr z jednoho rozdělení. Označme R_i^X pořadí veličin X_i ve spojeném výběru, uspořádaném podle velikosti a necht' $R^X = \sum_{i=1}^n R_i^X$. Potom je

$$E(R^X) = \frac{m(m+n+1)}{2}, \quad Var(R^X) = \frac{mn(m+n+1)}{12}$$

Test lze založit přímo na statistice R^X a kritický obor je potom určen nerovností $R^X \geq w_{m,n}(1 - \frac{\alpha}{2})$, kde kritické hodnoty $w_{m,n}(1 - \frac{\alpha}{2})$ jsou tabelovány (viz např. [An], [Sk], [Zv]). Pro přibližný test použijeme normovanou veličinu $W = \frac{R^X - E(R^X)}{\sqrt{D(R^X)}}$, která má za platnosti hypotézy pro velké rozsahy m a n přibližně rozdělení $N(0, 1)$.

Oba uvedené testy - znaménkový i Wilcoxonův - jsou častou používány jako testy **párové** namísto párového t-testu.

2.3 Regresní analýza

2.3.1 Regresní závislost

V matematice vyjadřujeme závislost hodnot jedné proměnné na hodnotách druhé proměnné funkčním vztahem. V praktických úlohách je však situace složitější. Při měření hodnot sledované veličiny, při jejíž realizaci působí řada dalších (náhodných) vlivů, dostáváme soubor naměřených hodnot, které vykazují často jisté odchylky proti hodnotám, které bychom očekávali

z teoretického rozboru sledovaného jevu nebo z jakési očekávané pravidelnosti.

Příklad 2.3.1 Při soustružení vzniká v místě obrábění na nástroji teplota, závislá na rychlosti posuvu nástroje. Mezi teplotou θ měřenou ve stupních Celsia a rychlostí posuvu v v metrech za minutu byl odvozen teoretický vztah $\theta = \alpha v^\beta$, kde α a β jsou konstanty, závislé na dalších podmínkách experimentu. Hodnoty, které byly naměřeny při laboratorním měření, však tomuto vztahu odpovídají jen velmi přibližně, jak lze vidět z grafu.

Předpokládejme, že sledovanou náhodnou veličinu Y lze vyjádřit jako funkci (zpravidla nenáhodných) veličin X_1, \dots, X_r a náhodné odchylky ϵ jako

$$Y = f(X_1, \dots, X_r; \theta_1, \dots, \theta_s) + \epsilon.$$

Funkce f se nazývá **regresní funkce** a $\theta_1, \dots, \theta_s$ nazýváme **parametry regrese**. O náhodné veličině ϵ , která se často nazývá neprávem „chybou“, předpokládáme, že má symetrické rozdělení se střední hodnotou 0 a rozptylem σ^2 . Obvyklý je předpoklad normálního rozdělení $N(0, \sigma^2)$. Uvedený vztah se nazývá **regresní model**. Podle druhu závislosti regresní funkce na neznámých parametrech $\theta_1, \dots, \theta_s$ potom hovoříme buď o *lineárním regresním modelu* nebo o *nelineárním regresním modelu*. Nadále se budeme zabývat pouze lineárním modelem.

Střední hodnota $E(Y)$ je potom funkcí hodnot veličin X_1, \dots, X_r a neznámých parametrů $\theta_1, \dots, \theta_s$. Tuto vlastnost vyjádříme vztahem

$$E(Y) = f(x_1, \dots, x_r; \theta_1, \dots, \theta_s),$$

kde x_1, \dots, x_r jsou naměřené hodnoty veličin X_1, \dots, X_r a $\theta_1, \dots, \theta_s$ jsou parametry.

Náhodné veličině Y se říká **vysvětlovaná proměnná**, veličinám X_1, \dots, X_r budeme říkat **vysvětlující proměnné**. Podle tvaru regresní funkce budeme mluvit o **přímkové, exponenciální, kvadratické, polynomické** a jiných regresích. V případě přímkové regrese rozlišujeme podle počtu vysvětlujících proměnných tzv. **jednoduchou regresi** s jednou vysvětlující proměnnou a **vícenásobnou regresi** s více vysvětlujícími proměnnými.

V zásadě zde máme dva problémy: určit tvar (typ) regresní funkce a Při vyšetřování regresní závislosti je regresní funkce zpravidla známa (vyplývá z teoretických vztahů) nebo se její tvar odhaduje (opticky, například podle X-Y grafu rozptýlenosti). Proto se v dalším textu omezíme na úlohu odhadu regresních parametrů předpokládané regresní funkce. K tomu nejčastěji používáme tzv. **metodu nejmenších čtverců**. Tato metoda spočívá v provedení n nezávislých měření

hodnot veličin Y, X_1, \dots, X_r a v nalezení hodnot $\hat{\theta}_1, \dots, \hat{\theta}_s$, při nichž funkce

$$S(\theta_1, \dots, \theta_s) = \sum_{i=1}^n \left[y_i - f(x_{1_i}, \dots, x_{r_i}; \theta_1, \dots, \theta_s) \right]^2$$

nabývá svého minima. Vektory $y_i, x_{1_i}, \dots, x_{r_i}$ označují i -té pozorování vektoru $Y, X_1, \dots, X_r, i = 1, \dots, n$.

Nejsme-li si jisti a rozhodujeme-li se mezi několika modely, potom zpravidla volíme ten, v němž je hodnota funkce $S(\theta_1, \dots, \theta_s)$ – takzvaný **reziduální součet čtverců** – nejmenší.

V případě lineární regresní funkce $f(x, \alpha, \beta) = \alpha + \beta x$ budeme minimalizovat funkci $S(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$. Nutnou podmínkou pro extrém funkce dvou proměnných je nulovost obou parciálních derivací

$$\begin{aligned} \frac{\partial S}{\partial \alpha} &= -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) = 0 \\ \frac{\partial S}{\partial \beta} &= -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) x_i = 0, \end{aligned}$$

což vede k takzvané **soustavě normálních rovnic**

$$\begin{aligned} n\alpha + \beta \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i \\ \alpha \sum_{i=1}^n x_i + \beta \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n y_i x_i \end{aligned}$$

jejímž řešením dostaneme bodové odhady a a b parametrů α a β

$$\begin{aligned} b &= \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ a &= \bar{y} - \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \bar{x} = \bar{y} - b\bar{x} \end{aligned}$$

kde $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. Podmínku postačující není třeba vyšetřovat, neboť funkce $S(\alpha, \beta)$ je ryze konvexní.

2.3.2 Jednoduchá přímková regrese

Velmi častým případem regresní závislosti je **přímková regrese**. Předpokládejme regresní vztah $Y = \alpha + \beta X + \epsilon$, kde X je náhodná veličina a ϵ je náhodná veličina s normálním rozdělením $N(0, \sigma^2)$.

Bodové odhady a a b parametrů α a β získáme metodou nejmenších čtverců ve tvaru uve-

deném v příkladě VIII.1.6. Naměřené hodnoty y_1, \dots, y_n lze považovat za hodnoty realizací nezávislých náhodných veličin Y_1, \dots, Y_n s normálním rozdělením $N(a + bx_i, \sigma^2)$. Z tohoto hlediska jsou bodové odhady a a b odhadovými statistikami, a tedy náhodnými veličinami. Hodnoty $e_i = y_i - a - bx_i, i = 1, \dots, n$ se nazývají **rezidua** a lze je považovat za odhady hodnot chybového členu ϵ . Číslo $\hat{y}_i = a + bx_i$ je odhadem hodnoty náhodné veličiny Y_i .

Označme $S_R = S(a, b)$ takzvaný **reziduální součet čtverců**

$$S_R = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 = \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n y_i - b \sum_{i=1}^n x_i y_i.$$

Bodový odhad s^2 rozptylu σ^2 chybového členu ϵ je potom dán vztahem $s^2 = \frac{S_R}{(n-2)}$ a nazývá se **reziduální rozptyl**.

Pomocí s^2 lze vyjádřit odhady rozptylu obou regresních parametrů

$$S_a^2 = \frac{s^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad S_b^2 = \frac{s^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$$

Statistiky $T_\alpha = \frac{(a-\alpha)}{S_a}$ a $T_\beta = \frac{(b-\beta)}{S_b}$ mají Studentovo t-rozdělení o $(n-2)$ stupních volnosti. Intervalové odhady pro parametry α a β jsou potom dány nerovnostmi

$$a - S_a t_{n-2}(1 - \frac{\gamma}{2}) \leq \alpha \leq a + S_a t_{n-2}(1 - \frac{\gamma}{2})$$

$$b - S_b t_{n-2}(1 - \frac{\gamma}{2}) \leq \beta \leq b + S_b t_{n-2}(1 - \frac{\gamma}{2})$$

kde $(1-\gamma)$ je koeficient spolehlivosti a $t_{n-2}(1-\frac{\gamma}{2})$ je $(1-\frac{\gamma}{2})$ -kvantil t -rozdělení o $(n-2)$ stupních volnosti.

V některých případech nás zajímá, zda hodnota některého z parametrů se liší významně od nulové hodnoty nebo ne a zda jej lze tudíž v regresní funkci vynechat. Oboustranné testy nulovosti regresních koeficientů lze založit na odhadových statistikách $T_a = \frac{a}{S_a}$ resp. $T_b = \frac{b}{S_b}$ a jim odpovídajícím kritickým oborům tak, že při splnění nerovnosti

$$|T_a| \geq t_{n-2}(1 - \frac{\gamma}{2}), \quad \text{resp.} \quad |T_b| \geq t_{n-2}(1 - \frac{\gamma}{2}),$$

zamítneme hypotézu o nulovosti parametru α , resp. β , na hladině významnosti γ .

Regresním modelem se snažíme vysvětlit změny - variabilitu - vysvětlované veličiny Y pomocí změn vysvětlující veličiny X . Podíl části variability Y vysvětlené modelem ku celkové variabilitě

Y , zpravidla vyjádřený v procentech, se nazývá **koeficient determinace** R^2 a je dán vztahy

$$R^2 = \frac{\sum_{i=1}^n (a + bx_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \left(1 - \frac{S_R}{\sum_{i=1}^n (y_i - \bar{y})^2} \right),$$

kde S_R je reziduální součet čtverců.

K úplné regresní analýze patří i **analýza reziduí**. Především by měly vyhovovat předpokladu normality, za kterého byly všechny předchozí výsledky odvozeny. Pokud tomu tak není, nelze výsledky považovat za důvěryhodné. K ověření shody hodnot reziduí s normálním rozdělením lze použít některý z testů, uvedených v odstavci VII.3, nebo pravděpodobnostní papír, který je popsán v kapitole X. Z analýzy reziduí lze detekovat i takzvaná **odlehlá pozorování**. To znamená ty hodnoty, které byly chybně naměřeny nebo indikují nesrovnalosti v modelu, a jimž je třeba věnovat zvláštní pozornost. Ke zjišťování těchto hodnot lze použít například krabicové gragy, popsané v kapitole X.

Model lineární regrese lze použít i v některých případech, kdy závislost mezi veličinami X a Y není lineární. Jsou to případy, kdy lze provést takzvanou **linearizaci modelu**. Vhodnou transformací převedeme nelineární závislost na lineární a použijeme lineární regresní model. Přitom však musíme být velmi opatrní, neboť vše, co bylo odvozeno pro lineární regresní model za předpokladu normality chybového členu ϵ platí pouze pro "linearizovaný model", nikoli pro model původní, a to opět za předpokladu, že náhodná veličina, odpovídající transformovanému chybovému členu v linearizovaném modelu, má normální rozdělení.

Příklad 2.3.2 *Závislost mezi teplotou θ a rychlostí posuvu v v příkladu 3.3.1. lze považovat za regresní závislost ve tvaru $\theta = \alpha \cdot v^\beta \cdot \epsilon$, kde α a β jsou regresní koeficienty a ϵ je náhodná veličina se střední hodnotou 1. Provedeme-li transformaci $Y = \ln\theta$, $X = \ln v$, $a = \ln\alpha$, $e = \ln\epsilon$, dostaneme $Y = a + bX + e$, tedy lineární vztah.*

Podobně jako v předchozím příkladu lze linearizovat i jiné modely, např. logaritmický, tj. $Y = \ln(\alpha + \beta \cdot X)$, reciproký $Y = \frac{1}{\alpha + \beta \cdot X}$ a další.

Kapitola 3

Jednofaktorové experimenty

Z předchozího textu je zřejmé, že velikost experiment se rychle zvyšuje s množstvím faktorů, nebo počtem jejich úrovní. Například, pokud máme sledovat dva faktory ve třech úrovních, je v případě úplného experimentu požadováno pozorování 9 (tedy 3×3) různých kombinací. Přidáme-li třetí faktor se třemi úrovněmi, počet kombinací naroste na 27 ($3 \times 3 \times 3$). Při čtvrtém faktoru se třemi úrovněmi už je to 81 kombinací. Použijeme-li pouze dvě úrovně pro každý faktor, pak v případě čtyř faktorů potřebujeme jen 16 ($2 \times 2 \times 2 \times 2$) kombinací. Z tohoto důvodu se často návrh experimentu omezuje na dvě úrovně. Dalšího snížení počtu pozorovaných kombinací lze dosáhnout při dílčích (neúplných) faktoriálních návrzích experimentů.

V této kapitole budeme uvažujeme situaci, kdy předpokládáme, že na odezvu působí pouze jediný faktor. Tento předpoklad je často zjednodušením reálného stavu, ale z praktického hlediska je poměrně dobře uchopitelný, realizovatelný a snadno pochopitelný.

Příklad 3.0.3 *Spotřeba nití v krejčovské výrobě se může lišit při použití různých typů šicích strojů. Uvažujme dva používané typy: stroje s asynchronním motorem a stroje se servomotorem. Je prokazatelný vliv typu šicího stroje na spotřebu?*

Příklad 3.0.4 *Pro tištění vodivých drah na textiliích nebo celkovou úpravu textilií vedoucí k elektrické vodivosti textilie se nabízí využití recyklovaných uhlíkových vláken v pružných kompozitních materiálech. Vzorčky kompozitních materiálů byly zhotoveny z nanočástic C-vláken v poměrech 50%, 70%, 80% a 90% s disperzním lepidlem. Je zkoumán vliv koncentrace uhlíkových vláken na povrchovou rezistivitu vzorků.*

Příklad 3.0.5 *Odolnost bakelitového krytu spínače závisí na teplotě, při které dochází k jeho vytváření. Jaká je optimální teplota pro dosažení nejlepší pevnosti?*

Příklad 3.0.6 Máme rozhodnout, která metoda je vhodnější pro měření zátrhovosti textilií. Testujeme 20 vzorků různých textilií a porovnáváme naměřené hodnoty jednotlivými metodami na každém vzorku.

Příklad 3.0.7 Předmětem experimentu je porovnání výparného odporu na různých druzích hašičských uniforem před a po praní. Pro měření bylo vybráno pět vzorků různých výrobců. Výsledkem jsou dvě hodnoty výparného odporu pro každý vzorek, které indikují vliv praní. Zajímá nás, který výrobce dodává výrobky vhodnější z hlediska komfortních vlastností, mezi něž výparný odpor patří.

Pět uvedených příkladů má jedno společné: vždy se jedná o předpokládanou závislost odezvy na jediném ovlivňujícím faktoru. Rozdíl je v tom, že v prvním případě faktor nabývá pouze dvou úrovní (dva typy šicích strojů), ve druhém případě rozlišujeme čtyři úrovně, zatímco ve třetím případě může faktor, kterým je teplota, nabývat teoreticky nekonečně mnoha hodnot (úrovní) z nějakého přípustného intervalu (daného technologickým postupem). Ve čtvrtém a pátém případě se jedná o pozorování "po dvojicích", kdy na jednom vzorku naměříme vždy dvě různé odezvy za různých úrovní ovlivňujícího faktoru. V těchto případech může hodnota odezvy záviset i na dalších vlastnostech jednotlivých vzorků.

Tyto rozdíly samozřejmě ovlivní způsob navržení experimentu a především analýzu jejich výsledků. V dalším textu si popíšeme některé základní přístupy k řešení podobných úloh.

3.1 Dvouúrovňové jednofaktorové experimenty

Uvažujme nyní situace, kdy faktor nabývá pouze dvou úrovní. V takovém případě potřebujeme provést měření či pozorování při každé z těchto úrovní. Dostaneme tak dvě skupiny hodnot odezvy, které mezi sebou budeme porovnávat.

Při návrhu takového experimentu musíme rozhodnout následující otázky:

- kolik budeme provádět replikací (měření)?
- jak vybereme jednotlivé případy?
- budeme sdružovat měření do bloků?
- jaké metody použijeme pro vyhodnocení výsledků?

Příklad 3.1.1 Kolikrát je třeba změřit fyzikální veličinu jejíž přesná hodnota je m , abychom mohli s pravděpodobností 0,96 tvrdit, že průměr těchto měření se liší od m o méně než 2? Je známo, že směrodatná odchylka měřicí metody je $\sigma = 4$.

Řešení: Přesná hodnota m představuje střední hodnotu měření $X_i, i = 1, \dots, n, DX_i = \sigma^2 = 16$. Pro $S_n = \sum_{i=1}^n$ je $ES_n = nm, DS_n = 16n$, použitím Čebyševovy nerovnosti (viz 2.1) a centrální limitní věty (viz ??) dostáváme

$$P\left(\left|\frac{S_n}{n} - m\right| < 2\right) = P\left(\left|\frac{S_n - nm}{4\sqrt{n}}\right| < \frac{\sqrt{n}}{2}\right) \approx 2\Phi\left(\frac{\sqrt{n}}{2}\right) - 1.$$

Poslední výraz má být roven alespoň 0,96. Úpravou této nerovnosti je

$$2\Phi\left(\frac{\sqrt{n}}{2}\right) \geq 0,96 \text{ a tedy } \frac{\sqrt{n}}{2} \geq \Phi^{-1}(0,96) = 2,055,$$

tedy počet měření musí být alespoň 17. Příklad představuje úlohu plánování experimentu na základě požadavků na přesnost.

3.1.1 Nezávislá měření

Jsou-li veličiny X a Y nezávislé, používáme pro srovnání středních hodnot dvou výběrů **dvouvýběrový t-test**. Nechť X_1, \dots, X_n je výběr z rozdělení $N(\mu_X, \sigma_X^2)$ a Y_1, \dots, Y_m je výběr z rozdělení $N(\mu_Y, \sigma_Y^2)$ a tyto výběry jsou na sobě nezávislé. Rozlišujeme dva případy:

- (1) Oba výběry mají **stejné rozptyly** $\sigma_X^2 = \sigma_Y^2$. Potom statistika

$$T = \frac{\bar{X} - \bar{Y}}{S} \sqrt{\frac{mn}{m+n}},$$

kde

$$S^2 = \frac{1}{m+n-2} [\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2]$$

má za platnosti nulové hypotézy $H : \mu_X = \mu_Y$ Studentovo t-rozdělení o $m+n-2$ stupních volnosti. Test uvedené hypotézy proti oboustranné alternativě $A : \mu_X \neq \mu_Y$ na hladině významnosti α lze tedy založit na nerovnosti

$$\frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} \geq t_{m+n-2}(1 - \frac{1}{2}\alpha).$$

Test hypotézy H proti jednostranným alternativám $A_1 : \mu_X \leq \mu_Y$, resp. $A_2 : \mu_X \geq \mu_Y$ na hladině významnosti α je založen na nerovnostech

$$\begin{aligned} \frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} &\geq t_{m+n-2}(1 - \alpha), \\ \frac{\bar{y} - \bar{x}}{S} \sqrt{\frac{mn}{m+n}} &\geq t_{m+n-2}(\alpha) = -t_{m+n-2}(\alpha)(1 - \alpha) \end{aligned}$$

(2) Oba výběry mají **různé rozptyly** $\sigma_X^2 \neq \sigma_Y^2$. Potom použijeme přibližný test, založený na statistice

$$\tilde{T} = \frac{\bar{X} - \bar{Y}}{\tilde{S}}, \text{ kde } \tilde{S}^2 = \frac{1}{n}s_X^2 + \frac{1}{m}s_Y^2.$$

Test hypotézy $H : \mu_X = \mu_Y$ proti oboustranné alternativě na hladině významnosti α lze založit na nerovnosti

$$\frac{|\bar{y} - \bar{x}|}{\tilde{S}} \geq \frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \frac{\alpha}{2}) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \frac{\alpha}{2}) \right]$$

na nerovnosti

$$\frac{\bar{y} - \bar{x}}{\tilde{S}} \geq \frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \alpha) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \alpha) \right]$$

při jednostranné alternativě $A_1 : \mu_X \leq \mu_Y$ a na

$$\frac{\bar{y} - \bar{x}}{\tilde{S}} \leq -\frac{1}{\tilde{S}^2} \left[\frac{1}{n}s_X^2 t_{n-1}(1 - \alpha) + \frac{1}{m}s_Y^2 t_{m-1}(1 - \alpha) \right]$$

při alternativě $A_2 : \mu_X \geq \mu_Y$.

Příklad 3.1.2 Při zpracování je třeba materiál zahřát na vysokou teplotu. Před zpracováním bylo vybráno náhodně 10 vzorků a změřena jejich tvrdost. Po zpracování bylo opět vybráno náhodně jiných 10 vzorků, na nichž byla změřena tvrdost. Naměřené hodnoty jsou v následující tabulce:

před	3,15	2,98	3,00	2,75	3,21	3,33	2,95	2,81	3,26	2,88
po	3,21	2,99	3,11	2,91	3,22	3,28	3,09	3,00	3,28	2,99

Testujte hypotézu, že se tvrdost materiálu vlivem zpracování nemění.

Řešení: Je $m = n = 10$. Spočteme $\bar{x} = 3,032$, $\bar{y} = 3,108$, $s_x^2 = 0,03875$, $s_y^2 = 0,018$. Za předpokladu, že rozptyl před i po zpracování zůstává stejný (naměřený rozdíl je nevýznamný), použijeme postup, popsáný v VII.2.6.a). Dostaneme hodnotu $s^2 = 0,02838$ a testové statistiky $T = 1,009$. Při hladině významnosti $\alpha = 0,05$ je $t_{18}(0,975) = 2,101$ a tedy hypotézu nelze zamítnout.

Příklad 3.1.3 Uvažujme stejnou úlohu jako v předchozím příkladu, pouze s tím rozdílem, že sledovaná veličina je měřena před i po zpracování na 10 vzorcích, které byly náhodně vybrány před začátkem experimentu. Naměřená data zůstávají stejná.

Řešení: V tomto případě je třeba vzít do úvahy závislost, která zde může být způsobena dalšími vlastnostmi vzorků. Proto použijeme párový t-test. Dostáváme $\bar{z} = \bar{x} - \bar{y} = -0,076$, $s_z = 0,07777$.

Testová statistika zde bude mít hodnotu $T = 3,09$, kterou budeme srovnávat s číslem $t_9 = (0,975) = 2,262$. V tomto případě hypotézu zamítneme. Uvedené příklady ukazují, jaký vliv na výsledek může mít tzv. *návrh experimentu*. Druhý případ lépe vystihuje skutečnost, že naměřená data nejsou nezávislá a bere do úvahy další možné vlivy, plynoucí z individuality vzorků.

Uvažujme dva nezávislé výběry: X_1, \dots, X_n z rozdělení $N(\mu_X, \sigma_X^2)$ a Y_1, \dots, Y_n z rozdělení $N(\mu_Y, \sigma_Y^2)$ můžeme provést tzv. **test shody rozptylů** neboli **F-test**. K testu hypotézy $H : \sigma_X^2 = \sigma_Y^2$ lze použít například statistiku

$$F = \frac{S_X^2}{S_Y^2}$$

Rozdělení statistiky F je podle V.4.6 a V.4.9 za předpokladu H rozdělením F o $(n-1)$ a $(m-1)$ stupních volnosti. Kritický obor pro test hypotézy H proti oboustranné alternativě $A : \sigma_X^2 \neq \sigma_Y^2$ na hladině významnosti α je určen nerovnostmi

$$\frac{s_X^2}{s_Y^2} \geq F_{n-1, m-1}(1 - \frac{\alpha}{2}) \text{ a } \frac{s_X^2}{s_Y^2} \leq F_{n-1, m-1}(\frac{\alpha}{2}) = \frac{1}{F_{n-1, m-1}(1 - \frac{\alpha}{2})}$$

kde $F_{n,m}(\alpha)$ je α -kvantil rozdělení $F(n, m)$. Tedy hypotézu zamítáme pro malá F blízká nule a pro velká F (při platnosti hypotézy by mělo být F blízké 1). Kritické obory pro test hypotézy H proti alternativám $A_1 : \sigma_X^2 \geq \sigma_Y^2$ a $A_2 : \sigma_X^2 \leq \sigma_Y^2$ na hladině významnosti α jsou určeny nerovnostmi

$$\frac{s_X^2}{s_Y^2} \geq F_{n-1, m-1}(1 - \alpha) \text{ resp. } \frac{s_X^2}{s_Y^2} \leq F_{n-1, m-1}(\alpha) = \frac{1}{F_{n-1, m-1}(1 - \alpha)}$$

V praxi se často používá tzv. *předpokladu normality*, t.j., že náhodný výběr pochází z normálního rozdělení s určitou střední hodnotou a nějakým, blíže neurčeným rozptylem. K ověření tohoto předpokladu lze použít **testy normality** založené na výběrových koeficientech **šikmosti** A_3 a **špičatosti** A_4 . Pro tyto statistiky (viz. V.3.4) platí následující vztahy:

$$E(A_3) = 0, \quad Var(A_3) = \frac{6(n-2)}{(n+1)(n+3)}$$

$$E(A_4) = -\frac{6}{n+1}, \quad Var(A_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

přičemž $\sqrt{n}A_3$ a $\sqrt{n}A_4$ mají při $n \rightarrow \infty$ přibližně normálního rozdělení. Test založený na šikmosti zamítně hypotézu o normalitě, pokud

$$\frac{|A_3|}{\sqrt{Var(A_3)}} \geq u(1 - \frac{\alpha}{2})$$

test založený na špičatosti zamítne hypotézu o normalitě, pokud

$$\frac{|A_4 - E(A_4)|}{\sqrt{\text{Var}(A_4)}} \geq u(1 - \frac{\alpha}{2})$$

kde $u(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$ -kvantil rozdělení $N(0, 1)$. Hladina významnosti obou testů je asymptoticky rovna α . Při ověřování normality je vhodné provést oba testy. Hypotézu nezamítneme teprve tehdy, pokud ji nelze zamítnout oběma testy zároveň. Tyto testy jsou v některých případech citlivější na porušení normality než χ^2 test, podrobnější informace nalezne čtenář v [An].

Příklad 3.1.4 Pomocí testů šikmosti a špičatosti testujte normalitu dat

z příkladu V.2.3.

Řešení: Pro výběrové koeficienty šikmosti a špičatosti dostáváme hodnoty $A_3 = -0,005364$, $E(A_3) = 0$, $D(A_3) = 0,129325$, $A_4 = -0,393762$, $E(A_4) = -0,146341$, $D(A_4) = 0,414962$. Testová statistika pro test založený na šikmosti nám dává hodnotu 0,015 a statistika pro test založený na špičatosti je rovna 0,384. Porovnáním s 0,975-kvantilem standardního normálního rozdělení $u(0,975) = 1,96$ tedy nelze zamítnout hypotézu o normalitě na hladině významnosti $\alpha = 0,05$ ani jedním z testů.

Dvouvýběrový Wilcoxonův test slouží k testování hypotézy o shodě distribučních funkcí dvou výběrů. Nechtě $\{X_1, \dots, X_n\}$ a $\{Y_1, \dots, Y_m\}$ jsou dva nezávislé výběry ze dvou spojitých rozdělení. Za platnosti hypotézy jsou tato rozdělení totožná a spojený výběr $\{X_1, \dots, X_n, Y_1, \dots, Y_m\}$ lze považovat za výběr z jednoho rozdělení. Označme R_i^X pořadí veličin X_i ve spojeném výběru, uspořádaném podle velikosti a nechtě $R^X = \sum_{i=1}^n R_i^X$. Potom je

$$E(R^X) = \frac{m(m+n+1)}{2}, \text{Var}(R^X) = \frac{mn(m+n+1)}{12}$$

Test lze založit přímo na statistice R^X a kritický obor je potom určen nerovností $R^X \geq w_{m,n}(1 - \frac{\alpha}{2})$, kde kritické hodnoty $w_{m,n}(1 - \frac{\alpha}{2})$ jsou tabelovány (viz např. [An], [Sk], [Zv]). Pro přibližný test použijeme normovanou veličinu $W = \frac{R^X - E(R^X)}{\sqrt{D(R^X)}}$, která má za platnosti hypotézy pro velké rozsahy m a n přibližně rozdělení $N(0, 1)$.

3.1.2 Závislá pozorování

Párový t-test. Sledujeme-li na jednom objektu dva podobné znaky zároveň, používáme náhodný výběr dvojic náhodných veličin $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$. O veličinách X_i a Y_i předpokládáme, že jsou **párově závislé**. Například měření vlastnosti materiálu před tepelným zpracováním a po něm na vybraných n vzorcích.

Předpokládejme, že $\{X_1, \dots, X_n\}$ je náhodný výběr z normálního rozdělení $N(\mu_X, \sigma^2)$ a Y_1, \dots, Y_n je náhodný výběr z rozdělení $N(\mu_Y, \sigma_Y^2)$. Veličiny X_i a Y_i mohou být párově závislé. Budeme testovat hypotézu o rovnosti středních hodnot $H : \mu_X = \mu_Y$ proti alternativě $A : \mu_X \neq \mu_Y$ na hladině významnosti α . V tomto případě budeme místo původně sledovaných veličin $(X_1, Y_1), \dots, (X_n, Y_n)$ pracovat s veličinami Z_1, \dots, Z_n , kde $Z_i = Y_i - X_i, i = 1, \dots, n$. Protože X_i a Y_i mají normální rozdělení, bude se i veličina Z řídit normálním rozdělením se střední hodnotou $\mu_Z = \mu_Y - \mu_X$ a rozptylem σ_Z^2 , o jehož vztahu k rozptylům σ_X a σ_Y nelze vzhledem k možné závislosti nic předpokládat. Rovnost středních hodnot X a Y je ekvivalentní nulovosti střední hodnoty rozdílu Z . Pro aritmetický průměr platí $\bar{z} = \bar{x} - \bar{y}$ a hodnotu výběrového rozptylu s_Z^2 spočteme podle vztahu

$$s_Z^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - y_i - \bar{x} + \bar{y})^2.$$

K testu hypotézy $H : \mu_Z = 0$ na hladině významnosti α použijeme jednovýběrový t-test (viz VII.2.2), tedy při oboustranné alternativě $A : \mu \neq 0$ hypotézu H zamítneme, pokud

$$\bar{z} = -\frac{s_Z}{\sqrt{n}} t_{n-1} \left(1 - \frac{\alpha}{2}\right) \text{ nebo } \frac{s_Z}{\sqrt{n}} t_{n-1} \left(1 - \frac{\alpha}{2}\right) < \bar{z}.$$

Další dva uvedené testy - znaménkový a Wilcoxonův - jsou častou používány jako testy párové namísto párového t-testu.

Znaménkový test je test o hodnotě mediánu. Předpokládáme, že výběr X_1, \dots, X_n je z rozdělení se spojitou distribuční funkcí $F(x)$. Budeme testovat hypotézu $H : \widetilde{x}_{50} = x_0$ proti alternativě $\widetilde{x}_{50} \neq x_0$ na hladině významnosti α . Vytvoříme posloupnost rozdílů $X_1 - x_0, \dots, X_n - x_0$. Označme Z počet členů této posloupnosti s kladným znaménkem a m počet nenulových rozdílů. Z je náhodná veličina s binomickým rozdělením (viz II.3.2) s parametry m a $\frac{1}{2}$. Pro malá m lze tedy stanovit kritický obor pro dané α určením celého čísla c tak, aby byly splněny nerovnosti

$$\sum_{i=1}^c \binom{m}{i} \left(\frac{1}{2}\right)^m \leq \frac{\alpha}{2} < \sum_{i=1}^{c+1} \binom{m}{i} \left(\frac{1}{2}\right)^m$$

Hypotézu H potom zamítneme, pokud $Z < c$ nebo $m - c < Z$. Pro větší m lze využít aproximace binomického rozdělení rozdělením normálním (viz věta v IV.2.1) a kritický obor vyjádřit pomocí $(1 - \frac{\alpha}{2})$ -kvantilu rozdělení $N(0, 1)$ nerovností $\frac{|2Z - m|}{\sqrt{m}} \geq u(1 - \frac{\alpha}{2})$. Pro jednostranné alternativy vytvoříme kritický obor analogicky jako v VII.2.1.

Jednovýběrový Wilcoxonův test. Předpokládáme, že $\{X_1, \dots, X_n\}$ je výběr z rozdělení se spojitou distribuční funkcí $F(x)$, která je symetrická kolem mediánu \widetilde{x}_{50} (neboli $F(\widetilde{x}_{50} - x) =$

$1 - F(\widetilde{x}_{50} + x)$. Budeme opět testovat hypotézu $H : (\widetilde{x}_{50} = x_0)$ proti alternativě $(\widetilde{x}_{50} \neq x_0)$ na hladině významnosti α . Podobně jako v VII.4.1 i v tomto případě vytvoříme posloupnost rozdílů $X_1 - x_0, \dots, X_n - x_0$ a dále budeme počítat pouze s nenulovými rozdíly, jejichž počet označíme m . Tuto posloupnost uspořádáme vzestupně podle absolutních hodnot a označíme R_i^+ pořadí náhodné veličiny $|X_i - x_0|$. Sečteme-li pořadí R_i^+ pro všechny členy, pro které je $X_i - x_0 > 0$ a tento součet označíme S^+ , dostaneme statistiku, pro kterou za platnosti hypotézy platí

$$E(S^+) = \frac{m(m+1)}{4}, \quad Var(S^+) = \frac{m(m+1)(2m+1)}{24}$$

a pro velká m je její rozdělení přibližně normální. Proto budeme pracovat raději s normovanou veličinou $V = \frac{S^+ - E(S^+)}{\sqrt{Var(S^+)}}$. Hypotézu tedy zamítneme, pokud $|V| \geq v(1 - \frac{\alpha}{2})$. Pro malé hodnoty m jsou kritické hodnoty $v(1 - \frac{\alpha}{2})$ tabelovány, pro velká m lze použít kvantily rozdělení $N(0, 1)$.

3.1.3 Uspořádání do bloků

Očekáváme-li, že se experimentální podmínky budou během zkoušek měnit (samozřejmě kromě cílených změn zkoumaného faktoru), je vhodné uspořádat zkoušky v blocích. V rámci každého bloku se, pokud je to možné, vystřídají všechny úrovně zkoumaného faktoru. Pořadí střídání úrovní faktoru je uvnitř každého bloku náhodné. Takovým blokům se říká úplné znáhodněné bloky. Není-li možné v rámci bloku vystřídát všechny úrovně zkoumaného faktoru (neúplné bloky), je plánování experimentu složitější. Tímto případem se zde ale nebudeme zabývat.

Příklad 3.1.5 *Pomocí experimentu zkoumáme vliv použitého katalyzátoru na výtěžek chemického procesu a chceme vybrat nejvhodnější ze čtyř typů. Odezvou je výtěžek chemického procesu, tj. množství vyráběné látky, zkoumaný faktor, katalyzátor, má čtyři úrovně. Předpokládejme, že můžeme provést dvacet čtyři zkoušek, při vyváženém návrhu to znamená vždy šest zkoušek se stejným katalyzátorem. Jedna várka vstupní suroviny ovšem stačí jen na provedení čtyř zkoušek.*

Při plánování experimentu je třeba počítat s tím, že kvalita vstupní suroviny bude v různých várkách kolísat a přispívat tak k větší experimentální chybě. Proto je vhodné volit uspořádání do bloků. Ty budou tvořeny jednotlivými várkami vstupní suroviny. Při jedné várce vstupní suroviny vystřídáme všechny úrovně zkoumaného faktoru (katalyzátoru). Jejich pořadí se pro každý blok volí náhodně. Výsledky zkoušek jsou uvedeny v tabulce a znázorněny bodovým diagramem.

Příklad 3.1.6 *Úkolem je zjistit, zdali 4 různé typy ostrí produkují různé záznamy při zjišťování tvrdosti materiálu. Zkušební stroj provede vryp do testovaného materiálu a podle hloubky vrypu se usuzuje na tvrdost materiálu.*

Úroveň	Várka					
	1	2	3	4	5	6
A1	87	79	82	89	83	78
A2	93	84	89	96	86	87
A3	88	80	84	91	83	82
A4	88	77	83	90	82	79

Tab. 3.1: Vliv použitého katalyzátoru

Zde existuje pouze jeden faktor – druh ostří. Hloubka vrypu může ovšem kromě typu ostří záviset do jisté míry i na materiálu, z něhož je vzorek zhotoven. Vzorky totiž nemusí být zcela homogenní, neboť mohou být různě tvrdé. Tím se do návrhu pokusu vnáší další vedlejší faktor, a to variabilita způsobená různými etalony ze zkušebního materiálu. Takže variabilita měření obsahuje ještě variabilitu způsobenou různými etalony.

Abychom tutu závislost eliminovali, rozhodneme se měřit na každém etalonu vrypy všech čtyř typů ostří. To znamená, že provedeme nejméně $4 \times 4 = 16$ pokusů. K tomu je zapotřebí tedy celkem 16 vzorků, čtyři z každého etalonu. Měření uspořádáme do čtyř bloků, kde blokem bude etalon, na němž se provedou vždy měření s každým typem ostří, tj. 4 vrypy. Protože v každém bloku budeme mít všechny typy ostří, hovoříme o úplném blokovém návrhu. Když pořadí pokusů v rámci každého bloku budeme provádět náhodně, máme návrh experimentu typu úplných znárodných bloků.

3.2 Jednofaktorové experimenty s více úrovněmi

3.2.1 Ordinální odezva

Pro větší počet uvažovaných úrovní působícího faktoru lze pro vyhodnocení jeho vlivu použít takzvané „jednofaktorové analýzy rozptylu“. Zkoumáme závislost odezvy Y na faktoru X , reprezentovaném diskrétními hodnotami (úrovněmi) x_1, x_2, \dots, x_k .

Jednofaktorová analýza rozptylu dává odpověď na otázku, zda lze určitým faktorem X vysvětlovat různost kvantitativní odezvy Y . Předpokládá provedení k nezávislých náhodných výběrů hodnot znaku Y o rozsazích n_1, n_2, \dots, n_k při úrovních x_1, x_2, \dots, x_k faktoru X . To jest celkem $n = \sum_{i=1}^k n_i$ měření. Základním předpokladem použití analýzy rozptylu je, že každý z k nezávislých výběrů znaku Y pochází z normálního rozdělení $N(\mu_i, \sigma^2)$ se stejným rozptylem σ^2 . Získáváme tedy celkem n pozorování náhodné veličiny Y rozdělených do k skupin, každá o rozsahu $n_i, i = 1, \dots, k$. V analýze rozptylu testujeme hypotézu $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ oproti alternativní hypotéze, že alespoň dvě hodnoty μ jsou různé.

Analýza rozptylu srovnává variabilitu uvnitř k skupin s variabilitou mezi skupinami, charak-

terizovanou variabilitou jejich středních hodnot. Jako míra variability zde slouží střední součty kvadratických odchylek uvnitř a mezi skupinami. Součet kvadratických odchylek uvnitř k -té skupiny je

$$SS_j = \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2, j = 1, \dots, k,$$

kde $\bar{y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$. Variabilita uvnitř skupin se potom podílí na celkovém součtu kvadratických odchylek v experimentu hodnotou

$$SS_E = \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2.$$

Označíme-li celkový průměr v modelu jako $\bar{y} = \frac{1}{n} \sum_{j=1}^k \sum_{i=1}^{n_j} y_{ij}$, potom součet kvadratických odchylek mezi skupinami se na celkovém součtu kvadratických odchylek v experimentu podílí touto částí

$$SS_F = \sum_{j=1}^k (\bar{y}_j - \bar{y})^2.$$

Platí tedy, že

$$SS_T = \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2 = SS_F + SS_E.$$

Tyto dvě části celkové variability mají různé „stupně volnosti“, tedy jsou složeny z různého počtu nezávislých sčítanců. SS_F má počet stupňů volnosti $k - 1$ (což odpovídá k skupinám pro jednotlivé úrovně faktoru), SS_E má $\sum_{j=1}^k (n_j - 1) = n - k$ stupňů volnosti ($n_j - 1$ v každé skupině). Odtud dostaneme tzv. „střední součty čtverců“

$$MSS_F = \frac{SS_F}{k - 1} \quad \text{a} \quad MSS_E = \frac{SS_E}{n - k}.$$

Pokud by platila hypotéza H_0 , měly by být všechny skupinové průměry přibližně stejné, tedy s téměř nulovou variabilitou mezi nimi. Potom by MSS_F mělo být zanedbatelné vzhledem k MSS_E . Budeme tedy testovat poměr

$$F = \frac{MSS_F}{MSS_E}.$$

Pokud jsou splněny předpoklady pro použití této metody, tedy pokud

- měřené veličiny jsou stochasticky nezávislé a
- měřené veličiny mají normální rozdělení se stejným rozptylem σ^2 ,

potom má veličina F známé rozdělení pravděpodobnosti, nazývané F -rozdělení s parametry $k - 1$

a $n - k$. To že je známé znamená, že umíme nalézt tzv. kritickou hodnotu $f(k - 1, n - k)$, kterou náhodná veličina s tímto rozdělením překročí jen s danou pravděpodobností α .

Rozhodovací pravidlo pro zamítnutí hypotézy H_0 je potom $F \geq f(k - 1, n - k)$.

3.2.2 Nominální odezva

Velmi často odezva nemá ordinální charakter, což znamená, že její uspořádání jejích hodnot (pokud je máme vyjádřeny číselnými hodnotami) nemá smysl. Například je-li odezvou pohlaví, barva látky, druh omaku a podobně.

Předpokládejme, že odezva může nabývat m hodnot, ovlivňovaných jedním kvalitativním faktorem, který může nabývat k různých úrovní, u nichž může ale nemusí mít uspořádání smysl. V takovém případě provádíme při n replikacích $N = n \cdot k \cdot m$ pozorování a budeme sledovat četnosti výsledků n_{ij} kdy při i -té úrovni faktoru nastala j -tá úroveň odezvy. Výsledky budeme zapisovat do takzvané *kontingenční tabulky*. Kontingenční tabulka je tabulka o k řádcích a m sloupcích

Příklad 3.2.1 Ovlivňuje barva očí Rh faktor?

Provedeme 400 pozorování, jejichž výsledky jsou v Tab. 3.2 a chceme testovat hypotézu o tom, že barva očí neovlivňuje Rh faktor.

barva očí	Rh^+	Rh^-	součet
modrá	35	65	100
hnědá	94	206	300
součet	129	271	400

Tab. 3.2: Zjištěné četnosti kombinací

Za předpokladu nezávislosti by (podle marginálních součtů) mělo platit:

barva očí	Rh^+	Rh^-	součet
modrá	32,25	67,75	100
hnědá	96,75	203,25	300
součet	129	271	400

Tab. 3.3: Odhadnuté četnosti kombinací

Pro test nezávislosti v kontingenční tabulce použijeme testovou statistiku ve tvaru

$$T = \sum \frac{(\text{pozorovaná četnost} - \text{očekávaná četnost})^2}{\text{očekávaná četnost}}$$

Testová statistika má chí-kvadrát rozdělení o $(k - 1) \times (m - 1)$ stupňů volnosti

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(n_{ij} - m_{ij})^2}{m_{ij}}$$

V našem příkladu je konkrétně

$$\chi^2 = \frac{(35 - 32,25)^2}{32,25} + \frac{(65 - 67,75)^2}{67,75} + \frac{(94 - 96,75)^2}{96,75} + \frac{(206 - 203,25)^2}{203,25} \doteq 0,2289$$

Tuto hodnotu budeme srovnávat s kritickou hodnotou chí-kvadrát rozdělení s jedním stupněm volnosti $\chi_1^2(0,05) = 3,84$. Tedy hypotézu o nezávislosti nezamítáme.

3.2.3 Latinské čtverce

Uvažujme modelovou situaci, kdy předpokládáme závislost hodnoty odezvy na jediném faktoru, který může mít celkem k úrovní. Nicméně, experiment nelze provést izolovaně od okolních podmínek, které mohou výsledek také ovlivnit. V tomto případě hovoříme o tzv. *vedlejších faktorech*. Pokud tyto vedlejší faktory mohou nabývat také k úrovní, můžeme pro návrh experimentu použít schéma, nazývané *latinský čtverec*.

Latinský čtverec je čtvercová tabulka $k \times k$, jejíž řádky odpovídají úrovním jednoho z vedlejších faktorů, řádky odpovídají hodnotám úrovní druhého z vedlejších faktorů. V buňkách tabulky jsou velkými latinskými písmeny¹ vyznačeny úrovně uvažovaného hlavního faktoru. Interpretace takové tabulky je následující: za působení každé z k úrovní vedlejšího faktoru provedeme k měření, v nichž se vystřídají všechny úrovně druhého z vedlejších faktorů a všechny úrovně hlavního faktoru. Cílem je provést k replikací měření s maximálním vyloučením vlivů vedlejších faktorů.

	1	2	3	4	5
1	A	B	C	D	E
2	E	A	B	C	D
3	D	E	A	B	C
4	C	D	E	A	B
5	B	C	D	E	A

Tab. 3.4: Příklad latinského čtverce 5×5 .

3.3 Regresní experimenty

V matematice vyjadřujeme závislost hodnot jedné proměnné na hodnotách druhé proměnné funkčním vztahem. V praktických úlohách je však situace složitější. Při měření hodnot sledované veličiny, při jejíž realizaci působí řada dalších (náhodných) vlivů, dostáváme soubor naměřených hodnot, které vykazují často jisté odchylky proti hodnotám, které bychom očekávali z teoretického rozboru sledovaného jevu nebo z jakési očekávané pravidelnosti.

¹Odtud název „latinský“ čtverec

Příklad 3.3.1 Při soustružení vzniká v místě obrábění na nástroji teplota, závislá na rychlosti posuvu nástroje. Mezi teplotou θ měřenou ve stupních Celsia a rychlostí posuvu v v metrech za minutu byl odvozen teoretický vztah $\theta = \alpha v^\beta$, kde α a β jsou konstanty, závislé na dalších podmínkách experimentu. Hodnoty, které byly naměřeny při laboratorním měření, však tomuto vztahu odpovídají jen velmi přibližně, jak lze vidět z grafu.

Předpokládejme, že sledovanou náhodnou veličinu Y lze vyjádřit jako funkci (zpravidla nenáhodných) veličin X_1, \dots, X_r a náhodné odchylky ϵ jako

$$Y = f(X_1, \dots, X_r; \theta_1, \dots, \theta_s) + \epsilon.$$

Funkce f se nazývá **regresní funkce** a $\theta_1, \dots, \theta_s$ nazýváme **parametry regrese**. O náhodné veličině ϵ , která se často nazývá neprávem „chybou“, předpokládáme, že má symetrické rozdělení se střední hodnotou 0 a rozptylem σ^2 . Obvyklý je předpoklad normálního rozdělení $N(0, \sigma^2)$. Uvedený vztah se nazývá **regresní model**. Podle druhu závislosti regresní funkce na neznámých parametrech $\theta_1, \dots, \theta_s$ potom hovoříme buď o *lineárním regresním modelu* nebo o *nelineárním regresním modelu*. Nadále se budeme zabývat pouze lineárním modelem.

Střední hodnota $E(Y)$ je potom funkcí hodnot veličin X_1, \dots, X_r a neznámých parametrů $\theta_1, \dots, \theta_s$. Tuto vlastnost vyjádříme vztahem

$$E(Y) = f(x_1, \dots, x_r; \theta_1, \dots, \theta_s),$$

kde x_1, \dots, x_r jsou naměřené hodnoty veličin X_1, \dots, X_r a $\theta_1, \dots, \theta_s$ jsou parametry.

Náhodné veličině Y se říká **vysvětlovaná proměnná**, veličinám X_1, \dots, X_r budeme říkat **vysvětlující proměnné**. Podle tvaru regresní funkce budeme mluvit o **přímkové, exponenciální, kvadratické, polynomické** a jiných regresích. V případě přímkové regrese rozlišujeme podle počtu vysvětlujících proměnných tzv. **jednoduchou regresi** s jednou vysvětlující proměnnou a **vícenásobnou regresi** s více vysvětlujícími proměnnými.

V zásadě zde máme dva problémy: určit tvar (typ) regresní funkce a Při vyšetřování regresní závislosti je regresní funkce zpravidla známa (vyplývá z teoretických vztahů) nebo se její tvar odhaduje (opticky, například podle X-Y grafu rozptýlenosti). Proto se v dalším textu omezíme na úlohu odhadu regresních parametrů předpokládané regresní funkce. K tomu nejčastěji používáme tzv. **metodu nejmenších čtverců**. Tato metoda spočívá v provedení n nezávislých měření

hodnot veličin Y, X_1, \dots, X_r a v nalezení hodnot $\hat{\theta}_1, \dots, \hat{\theta}_s$, při nichž funkce

$$S(\theta_1, \dots, \theta_s) = \sum_{i=1}^n \left[y_i - f(x_{1i}, \dots, x_{ri}; \theta_1, \dots, \theta_s) \right]^2$$

nabývá svého minima. Vektory $y_i, x_{1i}, \dots, x_{ri}$ označují i -té pozorování vektoru $Y, X_1, \dots, X_r, i = 1, \dots, n$.

Nejsme-li si jisti a rozhodujeme-li se mezi několika modely, potom zpravidla volíme ten, v němž je hodnota funkce $S(\theta_1, \dots, \theta_s)$ – takzvaný **reziduální součet čtverců** – nejmenší.

V případě lineární regresní funkce $f(x, \alpha, \beta) = \alpha + \beta x$ budeme minimalizovat funkci $S(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$. Nutnou podmínkou pro extrém funkce dvou proměnných je nulovost obou parciálních derivací

$$\begin{aligned} \frac{\partial S}{\partial \alpha} &= -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) = 0 \\ \frac{\partial S}{\partial \beta} &= -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) x_i = 0, \end{aligned}$$

což vede k takzvané **soustavě normálních rovnic**

$$\begin{aligned} n\alpha + \beta \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i \\ \alpha \sum_{i=1}^n x_i + \beta \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n y_i x_i \end{aligned}$$

jejímž řešením dostaneme bodové odhady a a b parametrů α a β

$$\begin{aligned} b &= \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ a &= \bar{y} - \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \bar{x} = \bar{y} - b\bar{x} \end{aligned}$$

kde $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. Podmínku postačující není třeba vyšetřovat, neboť funkce $S(\alpha, \beta)$ je ryze konvexní.

3.3.1 Jednoduchá přímková regrese

Velmi častým případem regresní závislosti je **přímková regrese**. Předpokládejme regresní vztah $Y = \alpha + \beta X + \epsilon$, kde X je náhodná veličina a ϵ je náhodná veličina s normálním rozdělením $N(0, \sigma^2)$.

Bodové odhady a a b parametrů α a β získáme metodou nejmenších čtverců ve tvaru uve-

deném v příkladě VIII.1.6. Naměřené hodnoty y_1, \dots, y_n lze považovat za hodnoty realizací nezávislých náhodných veličin Y_1, \dots, Y_n s normálním rozdělením $N(a + bx_i, \sigma^2)$. Z tohoto hlediska jsou bodové odhady a a b odhadovými statistikami, a tedy náhodnými veličinami. Hodnoty $e_i = y_i - a - bx_i, i = 1, \dots, n$ se nazývají **rezidua** a lze je považovat za odhady hodnot chybového členu ϵ . Číslo $\hat{y}_i = a + bx_i$ je odhadem hodnoty náhodné veličiny Y_i .

Označme $S_R = S(a, b)$ takzvaný **reziduální součet čtverců**

$$S_R = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 = \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n y_i - b \sum_{i=1}^n x_i y_i.$$

Bodový odhad s^2 rozptylu σ^2 chybového členu ϵ je potom dán vztahem $s^2 = \frac{S_R}{(n-2)}$ a nazývá se **reziduální rozptyl**.

Pomocí s^2 lze vyjádřit odhady rozptylu obou regresních parametrů

$$S_a^2 = \frac{s^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad S_b^2 = \frac{s^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$$

Statistiky $T_\alpha = \frac{(a-\alpha)}{S_a}$ a $T_\beta = \frac{(b-\beta)}{S_b}$ mají Studentovo t -rozdělení o $(n-2)$ stupních volnosti. Intervalové odhady pro parametry α a β jsou potom dány nerovnostmi

$$a - S_a t_{n-2}(1 - \frac{\gamma}{2}) \leq \alpha \leq a + S_a t_{n-2}(1 - \frac{\gamma}{2})$$

$$b - S_b t_{n-2}(1 - \frac{\gamma}{2}) \leq \beta \leq b + S_b t_{n-2}(1 - \frac{\gamma}{2})$$

kde $(1-\gamma)$ je koeficient spolehlivosti a $t_{n-2}(1 - \frac{\gamma}{2})$ je $(1 - \frac{\gamma}{2})$ -kvantil t -rozdělení o $(n-2)$ stupních volnosti.

V některých případech nás zajímá, zda hodnota některého z parametrů se liší významně od nulové hodnoty nebo ne a zda jej lze tudíž v regresní funkci vynechat. Oboustranné testy nulovosti regresních koeficientů lze založit na odhadových statistikách $T_a = \frac{a}{S_a}$ resp. $T_b = \frac{b}{S_b}$ a jim odpovídajícím kritickým oborům tak, že při splnění nerovnosti

$$|T_a| \geq t_{n-2}(1 - \frac{\gamma}{2}), \quad \text{resp.} \quad |T_b| \geq t_{n-2}(1 - \frac{\gamma}{2}),$$

zamítneme hypotézu o nulovosti parametru α , resp. β , na hladině významnosti γ .

Regresním modelem se snažíme vysvětlit změny - variabilitu - vysvětlované veličiny Y pomocí změn vysvětlující veličiny X . Podíl části variability Y vysvětlené modelem ku celkové variabilitě

Y , zpravidla vyjádřený v procentech, se nazývá **koeficient determinace** R^2 a je dán vztahy

$$R^2 = \frac{\sum_{i=1}^n (a + bx_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \left(1 - \frac{S_R}{\sum_{i=1}^n (y_i - \bar{y})^2} \right),$$

kde S_R je reziduální součet čtverců.

K úplné regresní analýze patří i **analýza reziduí**. Především by měly vyhovovat předpokladu normality, za kterého byly všechny předchozí výsledky odvozeny. Pokud tomu tak není, nelze výsledky považovat za důvěryhodné. K ověření shody hodnot reziduí s normálním rozdělením lze použít některý z testů, uvedených v odstavci VII.3, nebo pravděpodobnostní papír, který je popsán v kapitole X. Z analýzy reziduí lze detekovat i takzvaná **odlehlá pozorování**. To znamená ty hodnoty, které byly chybně naměřeny nebo indikují nesrovnalosti v modelu, a jimž je třeba věnovat zvláštní pozornost. Ke zjišťování těchto hodnot lze použít například krabicové gragy, popsané v kapitole X.

Model lineární regrese lze použít i v některých případech, kdy závislost mezi veličinami X a Y není lineární. Jsou to případy, kdy lze provést takzvanou **linearizaci modelu**. Vhodnou transformací převedeme nelineární závislost na lineární a použijeme lineární regresní model. Přitom však musíme být velmi opatrní, neboť vše, co bylo odvozeno pro lineární regresní model za předpokladu normality chybového členu ϵ platí pouze pro "linearizovaný model", nikoli pro model původní, a to opět za předpokladu, že náhodná veličina, odpovídající transformovanému chybovému členu v linearizovaném modelu, má normální rozdělení.

Příklad 3.3.2 *Závislost mezi teplotou θ a rychlostí posuvu v v příkladu 3.3.1. lze považovat za regresní závislost ve tvaru $\theta = \alpha \cdot v^\beta \cdot \epsilon$, kde α a β jsou regresní koeficienty a ϵ je náhodná veličina se střední hodnotou 1. Provedeme-li transformaci $Y = \ln\theta$, $X = \ln v$, $a = \ln\alpha$, $e = \ln\epsilon$, dostaneme $Y = a + bX + e$, tedy lineární vztah.*

Podobně jako v předchozím příkladu lze linearizovat i jiné modely, např. logaritmický, tj. $Y = \ln(\alpha + \beta \cdot X)$, reciproký $Y = \frac{1}{\alpha + \beta \cdot X}$ a další.

Kapitola 4

Vícefaktorové experimenty

Ve většině případů působí na odezvu více než pouze jediný faktor. Potom je třeba provést návrh vícefaktorového experimentu. Jeho návrh je komplikovanější, vyžaduje zpravidla více měření a omezení počtu sledovaných úrovní působících faktorů. Takto navržené experimenty jsou však mnohem efektivnější než experimenty zkoumající faktory po jednom v čase, které zahrnují studium změny jediného faktoru v danou chvíli a jejich vliv na výrobek nebo proces. Zatímco jednofaktorové experimenty jsou snadno pochopitelné, neumožňují vyšetřování toho, jak tento faktor ovlivňuje produkt nebo proces v přítomnosti dalších faktorů. Vztah, při kterém se mění vliv jednoho faktoru na výrobek nebo postup v závislosti na jednom nebo více dalších faktorech, se nazývá interakce. Často jsou vzájemné účinky faktorů důležitější než vliv jednotlivých faktorů. To proto, že v prostředí, v němž se výrobek nebo proces realizuje, působí mnoho faktorů dohromady a ne izolované výskyty jednotlivých faktorů v různých časech. Vezměme si příklad interakce mezi dvěma faktory při chemickém procesu, kde například teplota zvyšuje výtěžnost jen mírně a zároveň zvýšení tlaku samo o sobě nemá žádný vliv na výnos. Nicméně, pokud roste teplota i tlak současně, výnos rychle roste. Potom říkáme, že existuje interakce mezi těmito dvěma faktory, které ovlivňují chemickou reakci.

Metodika DOE zajišťuje, že všechny faktory a jejich interakce jsou systematicky zkoumány. Proto informace získané z analýzy DOE jsou mnohem spolehlivější a úplnější, než výsledky z jednofaktorového experimentu, které ignorují interakci a mohou vést k chybným závěrům.

Faktoriální experiment slouží ke zkoumání více faktorů najednou a umožňuje studovat neaditivní chování faktorů, tj. interakce. V úplném faktoriálním experimentu se vystřídají všechny možné kombinace úrovní zkoumaných faktorů, v dílčím faktoriálním experimentu jsou některé kombinace vynechány. Faktory zahrnuté do experimentu mohou mít dvě nebo více úrovní. Uvažujeme-li v experimentu dva faktory, jeden např. s dvěma a druhý se třemi úrovněmi, exis-

tuje šest kombinací úrovní těchto faktorů. Z označení experimentu 2x3 je vidět, kolik faktorů a s jakým počtem úrovní se zkoumá a kolik se provede zkoušek při jedné replikaci. Protože se počet zkoušek s rostoucím počtem úrovní faktorů rychle zvětšuje, je výhodné, hlavně v první fázi experimentování, kdy je počet N zařazených faktorů velký, volit u všech faktorů jen dvě úrovně. Experimenty potom označujeme $2N$. Kromě toho se při velkém počtu faktorů počet zkoušek zkracuje na polovinu, čtvrtinu atd. (dílicí faktoriální experiment). Někdy se u všech faktorů volí tři úrovně. Pokud je ovšem některý z faktorů kategoriální, použije se podle potřeby i vyšší počet úrovní. Faktoriální experimenty s faktory o dvou úrovních mají navíc velkou výhodu v jednoduchém vyhodnocení. Tyto experimenty se také v praxi nejčastěji používají.

V jedné replice experimentu se každá kombinace úrovní objeví jen jednou. Pro rozhodování o existenci vlivu faktorů je však lepší, provedou-li se alespoň dvě repliky experimentu. Rozhodneme-li se pro více replikací, dodržujeme přitom zásadu vyváženého návrhu, to znamená, že počet opakování je ve všech experimentálních bodech (tj. u všech kombinací) stejný. To také znamená, že různým úrovním jednoho faktoru odpovídá stejný počet výsledků. Výhody vyváženého návrhu byly již zmíněny v úvodu kapitoly 2. Podobně jako v případě jediného zkoumaného faktoru můžeme volit úplně znáhodněný návrh nebo uspořádání do bloků. Možnost snadného vyhodnocení výsledků oceníme v případě, kdy nemáme k dispozici statistický software, který by umožnil provést analýzu rozptylu pro více než dva faktory. Existují-li alespoň dvě repliky experimentu, lze při vyhodnocení použít t -test. V případě jedné repliky se doporučuje grafická metoda spočívající ve vynesení vypočtených efektů do pravděpodobnostního grafu. Máme-li statistický software, např. Statgraphics, Statisticu nebo Minitab, je nejjednodušší použít k vyhodnocení analýzu rozptylu. Pro zkoumání vlivu dvou faktorů postačí nástroj v Excelu. Případy se třemi a více faktory by se daly samozřejmě řešit s použitím příslušných vzorců, zápis se však stává poněkud nepřehledný a proto se zde, pokud půjde o analýzu rozptylu pro tři a více faktorů, omezíme jen na interpretaci výstupů z počítače

4.1 Dvoufaktorové experimenty

V této části budeme předpokládat, že na odezvu působí dva měřitelné faktory. Experiment uspořádáme tak, že zvolíme dvě úrovně u každého z faktorů a provedeme jedno měření za působení každé ze všech možných kombinací úrovní těchto faktorů. Označíme-li úrovně faktoru A jako A_1 a A_2 , úrovně faktoru B jako B_1 a B_2 , potom odpovídající měření budeme označovat y_{ij} , kde i a j jsou úrovně faktoru A a B . Využijeme-li všechny možné kombinace úrovní faktorů, můžeme návrh experimentu znázornit pomocí takzvané *matice návrhu experimentu*, která

vypadá takto (Tab. 4.1):

měření	<i>A</i>	<i>B</i>	<i>y</i>
1	–	–	y_{11}
2	+	–	y_{21}
3	–	+	y_{12}
4	+	+	y_{22}

Tab. 4.1: Designová matice dvoufaktorového návrhu se dvěma úrovněmi bez interakcí.

V tabulce znaménko minus (–) označuje nižší ze dvou úrovní, znaménko plus (+) potom vyšší ze dvou zvolených úrovní¹. Řádky matice odpovídají jednotlivým měřením v rámci experimentu. Pokud pro každou kombinaci úrovní faktorů provedeme více replikací měření, bude v každém řádku ve sloupci *y* tolik hodnot, kolik bude replikací.

To, co nás zpravidla v takovém experimentu zajímá, je vliv jednotlivých faktorů na velikost odezvy. Proto zavádíme pojem *efektu faktoru*:

Efekt faktoru = průměrná změna odezvy při změně úrovně faktoru

Velikost efektu vyjadřujeme jako velikost změny odezvy při změně úrovně odpovídajícího faktoru. V našem případě pro vyjádření velikosti efektu faktoru *A* spočteme průměrnou hodnotu odezvy při vyšší úrovni faktoru $\bar{y}_{A_2} = \frac{1}{2}(y_{21} + y_{22})$ a průměrnou hodnotu odezvy při nižší úrovni $\bar{y}_{A_1} = \frac{1}{2}(y_{11} + y_{12})$. Efekt faktoru *A* potom vyjádříme jako jejich rozdíl

$$\hat{A} = \bar{y}_{A_2} - \bar{y}_{A_1} = \frac{y_{21} + y_{22}}{2} - \frac{y_{11} + y_{12}}{2} = \frac{y_{21} + y_{22} - y_{11} - y_{12}}{2}.$$

Podobně pro efekt faktoru *B* dostáváme

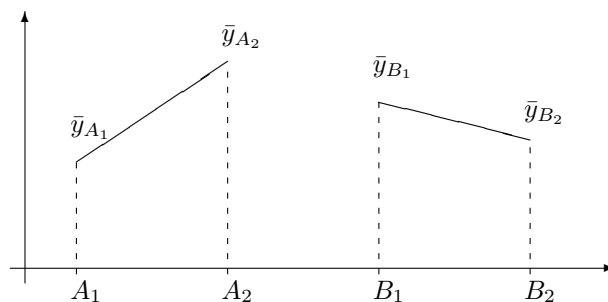
$$\hat{B} = \bar{y}_{B_2} - \bar{y}_{B_1} = \frac{y_{12} + y_{22}}{2} - \frac{y_{11} + y_{21}}{2} = \frac{y_{12} + y_{22} - y_{11} - y_{21}}{2}.$$

Bystrý čtenář si ihned všimne, že znaménka u y_{ij} ve zlomcích napravo odpovídají znaménkům v matici návrhu experimentu pro odpovídající faktory.

Srovnání vlivu jednotlivých faktorů lze provést také graficky. Na následujícím Obr. 4.1 jsou znázorněny průměrné hodnoty odezvy při různých úrovních faktorů *A* a *B*. Větší efekt má ten faktor, pro nějž má úsečka spojující tyto průměrné odezvy větší sklon (má v absolutní hodnotě vyšší směrnici). Na tomto obrázku je to faktor *A*.

Je-li rozdíl v hodnotě odezvy mezi dvěma úrovněmi jednoho faktoru výrazně jiný při různých úrovních druhého faktoru, např. $y_{21} - y_{11}$ a $y_{22} - y_{12}$, hovoříme o takzvané *interakci*:

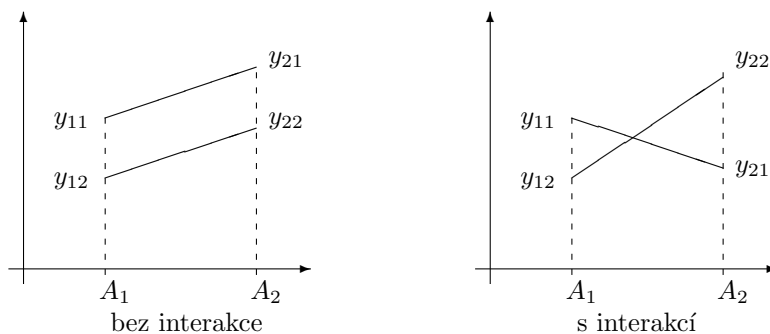
¹V některých případech se namísto znaménka – používá číslice 0 a namísto znaménka + číslice 1. To lze s výhodou použít v případě víceúrovňových experimentů, kdy nám už pouze dvě znaménka nestačí.



Obr. 4.1: Grafické znázornění vlivu jednotlivých faktorů.

Interakce mezi faktory = společné působení dvou či více faktorů na velikost odezvy

Na Obr. 4.2 vlevo je příklad situace, kdy efekty faktorů neinteragují, napravo je situace typická pro silnou interakce mezi faktory.



Obr. 4.2: Grafické znázornění vlivu jednotlivých faktorů.

V matici návrhu experimentu interakci označíme dvojicí písmen AB . Odpovídající znaménka pro interakci dostaneme ze znamének faktorů A a B tak, že tato znaménka spolu „vynásobíme“, tedy dvojice stejných znamének dá znaménko plus, dvojice různých znamének dá znaménko minus (viz Tab. 4.2). Interpretace tohoto pravidla může být například taková, že pokud interakce existuje, její vyšší vliv bude při stejné úrovni (nízké nebo vysoké) obou faktorů. Naopak, nízká úroveň interakce bude očekávána při rozdílných úrovních faktorů A a B .

měření	A	B	AB	y
1	-	-	+	y_{11}
2	+	-	-	y_{21}
3	-	+	-	y_{12}
4	+	+	+	y_{22}

Tab. 4.2: Designová matice úplného dvoufaktorového návrhu se dvěma úrovněmi.

Nyní můžeme vyjádřit velikost efektu interakce AB podobně jako efekty pro jednotlivé faktory: je to rozdíl průměrné odezvy při společném působení obou faktorů $\bar{y}_{AB+} = \frac{1}{2}(y_{11} + y_{22})$ a

průměrné odezvy při rozdílném působení obou faktorů $\bar{y}_{AB-} = \frac{1}{2}(y_{21} + y_{12})$, tedy

$$\widehat{AB} = \bar{y}_{AB+} - \bar{y}_{AB-} = \frac{y_{11} + y_{22}}{2} - \frac{y_{21} + y_{12}}{2} = \frac{y_{11} + y_{22} - y_{21} - y_{12}}{2}.$$

Cvičení: Uvažujme experiment, v němž byly provedeny dvě replikace pro každé měření. To znamená, že bylo provedeno celkem $2 \times 2 \times 2$ měření. Odvoďte vzorce pro efekty jednotlivých faktorů a interakce.

Řešení: Experimentu odpovídá následující matice návrhu:

měření	A	B	AB	y
1	-	-	+	$y_{11,1}, y_{11,2}$
2	+	-	-	$y_{21,1}, y_{21,2}$
3	-	+	-	$y_{12,1}, y_{12,2}$
4	+	+	+	$y_{22,1}, y_{22,2}$

Tab. 4.3: Designová matice úplného dvoufaktorového návrhu se dvěma úrovněmi při dvou replikacích.

Odpovídající efekty spočteme podle následujících vztahů (s využitím znamének z tabulky):

$$\begin{aligned} \hat{A} &= \bar{y}_{A_2} - \bar{y}_{A_1} = \frac{y_{21,1} + y_{21,2} + y_{22,1} + y_{22,2} - y_{11,1} - y_{11,2} - y_{12,1} - y_{12,2}}{4}, \\ \hat{B} &= \bar{y}_{B_2} - \bar{y}_{B_1} = \frac{y_{12,1} + y_{12,2} + y_{22,1} + y_{22,2} - y_{11,1} - y_{11,2} - y_{21,1} - y_{21,2}}{4}, \\ \widehat{AB} &= \bar{y}_{AB+} - \bar{y}_{AB-} = \frac{y_{11,1} + y_{11,2} + y_{22,1} + y_{22,2} - y_{21,1} - y_{21,2} - y_{12,1} - y_{12,2}}{4}. \end{aligned}$$

4.2 Vyhodnocení vícefaktoriálních experimentů

Statistický model pro dva faktory A, B:

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

kde

y_{ij} je odezva,

μ je společná úroveň pro oba faktory,

α_i označuje vliv i -té úrovně faktoru A ,

β_j označuje vliv j -té úrovně faktoru B ,

$(\alpha\beta)_{ij}$ je vliv interakce i -té a j -té úrovně faktorů A a B ,

ϵ_{ij} představuje náhodnou chybu při i -té a j -té úrovni faktorů A a B .

4.3 Vícefaktoriální experimenty

Statistický model pro tři faktory A, B a C:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijk}$$

kde

y_{ijk} je odezva,

μ je společná úroveň pro oba faktory,

$\alpha_i, \beta_j, \gamma_k$ označuje vlivy úrovní i, j, k faktorů A, B, C,

$(\alpha\beta)_{ij}, (\alpha\gamma)_{ik}, (\beta\gamma)_{jk}$ označuje vlivy interakcí druhého řádu, tedy interakcí dvojic faktorů AB, AC, BC při odpovídajících úrovních i, j, k ,

$(\alpha\beta\gamma)_{ijk}$ označuje vlivy interakce třetího řádu, tedy interakce všech tří faktorů dohromady při odpovídajících úrovních i, j, k ,

ϵ_{ij} představuje náhodnou chybu při i -té a j -té úrovni faktorů A a B.

test	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	-	-	+	+	-	-	+
5	+	+	-	+	-	-	-
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

Tabulka 4.1: Designová matice úplného třífaktorového návrhu se dvěma úrovněmi

4.4 Směšování ve vícefaktoriálních návrzích

S rostoucím počtem faktorů vzrůstá v úplném faktoriálním experimentu rychle počet zkoušek. Proto se často z úsporných důvodů provádí dílčí faktoriální experiment. Ukazuje se totiž, že v mnoha případech lze požadovanou informaci získat na základě vhodně vybrané části úplného faktoriálního experimentu. Experiment 2^N pro N faktorů s dvěma úrovněmi lze zkrátit na polovinu, čtvrtinu atd. původního rozsahu. Značíme jej potom $2^{N-1}, 2^{N-2}$ atd. Zkoumáme stále všech N faktorů, do experimentu však zahrneme pouze některé kombinace úrovní faktorů, takže celkový počet zkoušek bude $2^N/2, 2^N/4$ atd. Zkrácení experimentu znamená ztrátu části informace. To se projeví při vyhodnocení výsledků. Po vynechání části zkoušek nebudeme moci odhadnout velikost všech efektů, které by přicházely v úvahu v úplném faktoriálním experimentu. Avšak

vzhledem k tomu, že některé ze zkoumaných faktorů se nakonec ukáží jako nedůležité a že interakce vyššího řádu bývají zanedbatelné, tato skutečnost příliš nevádí. Zkracovat lze i experimenty s faktory o více úrovních, poměrně časté, zejména u Taguchiho přístupu, jsou dílčí faktoriální experimenty s faktory o třech úrovních. Značí se 3^{N-1} , 3^{N-2} atd. V těchto skriptech se však omezíme jen na dílčí faktoriální experimenty s dvěma úrovněmi faktorů, které jsou nejběžnější. Následující schéma představuje ukázkou zkrácení na šestnáctinu zkoušek. Místo $2^7 = 128$ zkoušek se provede jen 8 zkoušek např. při vyznačených kombinacích.

Výběr kombinací se tedy neprovádí náhodně, ale tak, aby důsledky zkrácení experimentu z hlediska účinnosti vyhodnocení byly co nejmenší. Po vynechání části kombinací úrovní faktorů zůstane zachována podstatná vlastnost, ortogonalita, která zjednodušuje analýzu výsledků a usnadňuje interpretaci. V literatuře zabývající se navrhováním experimentů i v příslušném softwaru jsou všechny návrhy zkrácených faktoriálních experimentů k dispozici a uživatel tedy nemusí kombinace úrovní faktorů sám určovat. Při výběru vhodného návrhu a stupně zkrácení experimentu je ovšem třeba rozumět určitým schématům a zápisům a proto zde bude na několika nejjednodušších příkladech podstata zkracování vysvětlena.

Uvažujme návrh 2^3 s 8 kombinacemi úrovní faktorů (úplný návrh). Z nějakých důvodů je třeba snížit počet měření na polovinu - vytvoříme "dílčí" faktoriální experiment 2^{3-1} . Podobně jako při dělení na bloky, použijeme k tomu poslední sloupec ABC (viz Tab. 4.4.2). Interakci ABC budeme nazývat *generátorem* plánu.

test	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
a	+	-	-	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
ab	+	+	-	+	-	-	-
ac	+	-	+	-	+	-	-
bc	-	+	+	-	-	+	-
abc	+	+	+	+	+	+	+

Tabulka 4.2: Designová matice úplného třífaktorového návrhu se dvěma úrovněmi

Experiment zkrácený na polovinu označujeme 2^{N-1} , kde N je počet faktorů. V případě tří zkoumaných faktorů ($N = 3$) to znamená, že ze všech možných kombinací úrovní faktorů A, B, C , kterých je osm, vybíráme čtyři. Při zkracování experimentu na polovinu vycházíme ze znamének interakce nejvyššího řádu, v našem případě ze znamének interakce ABC . Jak je vidět v tabulce, v polovině případů má interakce ABC kladné znaménko, v polovině případů záporné znaménko. Pro experiment 2^{3-1} vybereme čtyři kombinace úrovní vedoucí ke stejnému znaménku ABC , tedy buď kombinace ve stínovaných rádcích, kdy má interakce ABC znaménko $+$ nebo naopak

kombinace ve zbylých řádcích, kdy má interakce ABC znaménko $-$.

Zavedme pro sloupce designové matice operaci „ \cdot “ jako logické „and“, pro které platí

$$+ \cdot - = -, \quad + \cdot + = +, \quad - \cdot - = +,$$

potom zřejmě platí

$$A \cdot B = AB, \quad A \cdot C = AC, \quad B \cdot C = BC, \quad A \cdot B \cdot C = ABC.$$

Označme I značí sloupec samých „+“. Potom zřejmě

$$A \cdot A = A^2 = B \cdot B = B^2 = C \cdot C = C^2 = I, \quad A \cdot I = A, \quad B \cdot I = B, \quad C \cdot I = C.$$

Podmínku výběru potom můžeme označit symbolicky $ABC = I$. „Vynásobením“ této rovnosti sloupcem A, B nebo C dostáváme

$$\begin{aligned} A(ABC) &= IA, & B(ABC) &= BI, & C(ABC) &= CI, \\ IBC &= A, & IAC &= B, & ABI &= C, \\ BC &= A, & AC &= B, & AB &= C. \end{aligned}$$

Jak tyto rovnosti interpretovat nám ukáž pohled na následující tabulku.

test	A	B	C	AB	AC	BC	ABC	y
a	+	-	-	-	-	+	+	y_1
b	-	+	-	-	+	-	+	y_2
c	-	-	+	+	-	-	+	y_3
abc	+	+	+	+	+	+	+	y_4

Tabulka 4.3: Designová matice dílčího třífaktorového návrhu se dvěma úrovněmi

Protože jsou ve sloupci ABC dílčího xperimentu všechna znaménka stejná, nemůžeme odhadnout efekt interakce ABC . Porovnáme-li znaménka ve sloupcích A a BC , zjistíme, že sloupce jsou totožné. Znamená to, že pomocí vztahu

$$\frac{y_1 + y_4}{2} - \frac{y_2 + y_3}{2}$$

kde y_1 až y_4 jsou hodnoty odezvy naměřené při daných kombinacích úrovní, určujeme vlastně součet efektu faktoru A a efektu interakce BC , tedy $A + BC$. Z tohoto důvodu nepoznáme, zda je

rozdíly mezi průměry výsledků způsoben změnou úrovně faktoru A nebo změnami úrovní faktorů B a C . Tomuto jevu říkáme směřování efektů a efekty faktoru A a interakce BC označujeme někdy jako alias efekty. K podobnému zjištění dojdeme porovnáním sloupců B a AC nebo C a AB . Hlavní efekt faktoru B je smíšen s efektem interakce AC a hlavní efekt C s interakcí AB .

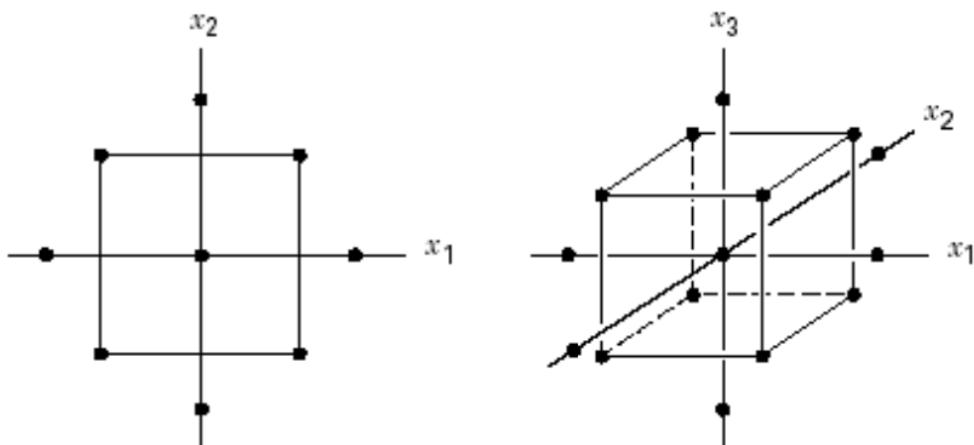
4.5 Významné body experimentu

Při návrhu experimentu rozlišujeme tři druhy takzvaných významných bodů. Jsou to vlastně navržená měření, lišící se v nastavení různých úrovní faktorů. V případě kdy každý z faktorů bude v experimentu vystupovat ve dvou úrovních, můžeme tyto označit hodnotami $-1 =$ nižší úroveň, $+1 =$ vyšší úroveň. Přiřadíme nyní každému měření vektor, jehož jednotlivé složky budou odpovídat faktorům a jejich hodnoty úrovním těchto faktorů.

Významné body návrhu:

- krychlové body
- centrální body
- hvězdicové body

Krychlové body: jsou v plánu vždy a je jich 2^{k-p} . Mají souřadnice ± 1 . Sledujeme-li například vliv tří faktorů, dostáváme 8 bodů o souřadnicích $(\pm 1, \pm 1, \pm 1)$. Nakreslíme-li tyto body do kartézské soustavy souřadnic, dostaneme vrcholy krychle se středem v počátku a délkou strany 2 (viz obr. Obr. 4.3). Krychlové body slouží k výpočtu efektů faktorů nebo koeficientů v lineárním modelu.



Obr. 4.3: Významné body experimentu pro dva faktory (vlevo) a pro tři faktory (napravo)

Centrální body: se doplňují do návrhu experimentu dodatečně a zpravidla se používají k odhadu rozptylu σ^2 . To je užitečné zvláště v případě, kdy měření odpovídající centrálním bodům nejsou replikována, nebo počet jejich replikací je velmi malý a nelze z nich odhadnout rozptyl σ^2 . Doporučovaný počet těchto bodů je 3–5). Centrální body mají v našem zobrazení nulové souřadnice (pro 3 faktory je centrální bod $(0, 0, 0)$). To odpovídá další úrovni sledovaných faktorů, která je někde mezi dolní a horní uvažovanou úrovní (nemusí být nutně uprostřed). Centrální body se nepoužívají k výpočtu efektů faktorů, ale z reziduí v centrálních bodech můžeme spočítat čistou chybu měření.

Hvězdicové body: Pro odhad koeficientů v lineárním modelu použijeme výsledky měření, odpovídající krychlovým bodům. Pro odhad koeficientů v úplném kvadratickém modelu však potřebujeme další měření. Ta jsou reprezentována hvězdicovými body. Nejlépe si je můžeme znázornit na trojrozměrném obrázku xxx. Představme si kulovou plochu se středem v počátku a procházející vrcholy krychle. Hvězdicové body leží na průsečících této kulové plochy se souřadnými osami. V trojrozměrném případě je jejich počet 6 (2 na každé ose), obecně jich je $2k$. Souřadnice pro tři faktory jsou $(\pm\alpha, 0, 0)$, $(0, \pm\alpha, 0)$, $(0, 0, \pm\alpha)$. Symbol $\pm\alpha$ zde charakterizuje další úroveň sledovaných faktorů, které jsou pod nižší nebo nad vyšší úrovněmi. Hvězdicové body navíc ještě zpřesňují odhady regresních koeficientů.

Kapitola 5

Taguchiho robustní návrhy

Robustní návrh je další možností, jak zlepšit jakost průmyslových procesů. Hlavní myšlenka spočívá v redukci variability procesu nikoli přímo pomocí náročné regulace veličin, které ji způsobují, ale vhodným nastavením jiných, snadněji ovladatelných faktorů. Variabilitu procesu způsobuje jednak proměnlivost různých měřitelných veličin, jako jsou např. změny teploty olejové lázně během času při zušlechťování oceli, jednak rozdíly mezi jednotkami vyráběnými souběžně na několika strojích, na několika pozicích apod., např. rozdíly mezi destičkami umístěnými na různých stranách šestibokého hranolu při výrobě integrovaných obvodů. V normálním procesu může být udržení teploty lázně na požadované neměnné úrovni příliš náročné nebo dokonce nemožné. Zjistíme-li však, že na variabilitu průhybu pružiny způsobenou kolísáním teploty má vliv doba zahřívání pružin během tepelného zpracování, místo náročné regulace teploty se zaměříme na zajištění vhodné doby zahřívání. Variabilitu jednotek vyráběných na několika strojích můžeme snížit pečlivým seřízením strojů, avšak odstranění rozdílů mezi jednotkami odebíranými z různých pozic šestibokého hranolu ve zmíněném příkladu se zdá být při stávající konstrukci zařízení nemožné. Ověříme-li, že na variabilitu tloušťky vrstvy vytvářené na povrchu křemíkových destiček v různých pozicích hranolu, na němž jsou umístěny, má vliv poloha trysky, z níž proudí chemická pára, soustředíme se na vyhledání optimální polohy této trysky. Uvedené příklady naznačují, že v popředí zájmu je variabilita odezvy, kterou se snažíme snížit. Cílem robustního návrhu je snížit citlivost sledované veličiny na kolísání rušivých veličin.

Popsaný problém lze řešit různými způsoby. Existují rozdíly jak v přípravě experimentu, tak ve vyhodnocení výsledků experimentu. Abychom mohli používané postupy popsat, rozšíříme dosud používanou terminologii o další pojmy. V robustním návrhu se rozlišují dvě hlavní skupiny faktorů. Řiditelné faktory jsou ovladatelné jak v normálním procesu, tak během experimentování, to znamená, že jejich úroveň je možné po nastavení udržet neměnné. Hodnoty rušivých faktorů se

během normálního procesu obvykle mění v čase, případně s polohou, během experimentu je však můžeme, aspoň do jisté míry, udržet na konstantní hodnotě. Za rušivé faktory považujeme i kategoriální faktory, např. faktor představující pozici v peci. Společnou vlastností měřitelných i kategoriálních rušivých faktorů je to, že existence různých úrovní během normálního procesu je zdrojem nežádoucí variability sledovaného znaku. Předpokládá se, že chování měřitelných rušivých veličin v normálním procesu aspoň přibližně známe. Jsou-li rušivé faktory měřitelnými veličinami, volí se obvykle dvě jejich úrovně, např. ve vzdálenosti směrodatná odchylka od předpokládané střední hodnoty. Do experimentu jsou tedy zahrnuty jak říditelné, tak rušivé faktory.

Myšlenka robustního návrhu má původ v Japonsku a jejím autorem je Taguchi. Princip robustního návrhu je všeobecně uznáván, Taguchiho metody analýzy experimentálních výsledků již méně. Z pohledu statistiků jde často o teoreticky nepodložené postupy nebo o postupy málo účinné, to znamená, že k rozpoznání vlivu faktoru je třeba provést větší počet zkoušek. Metody však pro svou zdánlivou jednoduchost získaly v praxi značnou oblibu a staly se součástí některých statistických paketů obsahujících modul pro navrhování experimentů (např. Minitab). Vymezený rozsah skript neumožňuje zabývat se danou problematikou v celé šíři. Omezíme se zde proto na návrh experimentu používaný v originálním přístupu, ukážeme zde však nejen Taguchiho způsob vyhodnocení, ale také jednu z dalších možností.

Optimalizace

Jakmile je pozornost zúžena na důležité faktory ovlivňující proces, dalším krokem je určit nejlepší nastavení těchto faktorů k dosažení požadovaného cíle. V závislosti na produktu nebo procesu, který je předmětem šetření tohoto cíle může být buď maximalizovat, minimalizovat nebo dosažení cílové hodnoty odezvy.

Testování robustnosti

Jakmile je hotovo optimální nastavení faktorů, které byly stanoveny, je důležité zjistit citlivost produktu nebo procesu na změny, které by mohly nastat v oblasti aplikačního prostředí. Tyto změny vyplývají ze změn faktorů, které mohou ovlivnit proces, ale jsou mimo kontrolu analytika. Takové faktory jako vlhkost, teplota okolí, změny v materiálu, apod. se označují jako šumové (vedlejší) faktory. Je důležité identifikovat zdroje takových změn a přijmout opatření, která by eliminovala jejich vlivy (například rozdělením experimentu do bloků).

5.1 Designové matice

Taguchiho experimenty vycházejí z tzv. ortogonálních oblastí či seskupení, které tvoří složené designové matice. Obvykle se tyto ortogonální oblasti označují písmenem L_k , kde index k znamená celkový počet pozorování v rámci experimentu. Nejzákladnějšími oblastmi jsou L_4 , L_8 , L_{16} , a L_{32} . Vlastně se jedná o dílčí faktoriální návrhy typu 2^{k-1} , kde $k = 3; 4; 5$ a 6 , kde jsou pouze jinak seřazené sloupce a namísto označení pro verze faktorů „+“ či „-“ je použito číslic 1 a 2.

L_4	A	B	C
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

Tab. 5.1: Designová matice návrhu L_4

L_8	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

Tab. 5.2: Designová matice návrhu L_8

Pro Taguchiho designové matice se používá označení $L_n(X^y)$, kde

n = počet řádek v návrhu,

X = počet úrovní faktorů,

y = počet sloupců v návrhu = počet faktorů.

Mezi základní typy designových matic v Taguchiho přístupu patří

- $L_4(2^3)$ – dvouúrovňový návrh pro 2–3 faktory
- $L_8(2^7)$ – dvouúrovňový návrh pro 3–7 faktorů
- $L_{12}(2^{11})$ – dvouúrovňový návrh pro 4–11 faktorů
- $L_{16}(2^{15})$ – dvouúrovňový návrh pro 4–15 faktorů
- $L_9(3^4)$ – tříúrovňový návrh pro 2–4 faktory
- $L_{18}(2^13^7)$ – návrh pro 1 faktor se 2 úrovněmi a 3–7 faktorů o třech úrovních
- $L_{16}(4^5)$ – čtyřúrovňový návrh pro 2–5 faktorů

5.2 Řízené a šumové faktory

V původním přístupu se pro každou skupinu faktorů uvažuje zvláštní návrh, tzv. vnitřní a vnější pole. Vnitřní pole je tvořeno kombinacemi říditelných faktorů, vnější pole rušivými faktory.

- říditelné (řízené) faktory jsou takové, u nichž jsme schopni nastavit jejich úroveň. Řízené faktory tvoří tzv. „vnitřní ortogonální sestavu“
- rušivé (neřízené, šumové) faktory lze pouze měřit. Nejsme schopni nastavit jejich úroveň, ani je nijak ovlivňovat. Tyto faktory tvoří tzv. „vnější ortogonální sestavu“.

Kombinace úrovní faktorů odpovídají faktoriálnímu či dílčímu faktoriálnímu návrhu. Obě pole jsou křížena, což znamená, že v každém bodě vnitřního pole, tj. při dané kombinaci úrovní říditelných faktorů, se vystřídají všechny body vnějšího pole, tj. kombinace úrovní rušivých faktorů. Provádí se tolik replikací, kolik je kombinací úrovní ve vnější sestavě. V každém bodě vnitřního pole se z naměřených hodnot sledované veličiny vypočtou určité charakteristiky a dále jsou analyzovány hodnoty těchto souhrnných charakteristik. Jednou z charakteristik je průměr, druhá charakteristika měří variabilitu hodnot odezvy. Taguchi používá charakteristiku „signál/šum“

$$S/N = \frac{\text{síla signálu}}{\text{síla šumu}} = \frac{(\text{citlivost})^2}{(\text{variabilita})^2} = \frac{\mu^2}{\sigma^2}.$$

Charakteristik typu signál/šum existuje celá řada, nejběžnější jsou však následující čtyři, používané při statickém návrhu, kdy sledovaná veličina (výstup procesu) by měla být konstantní (na rozdíl od dynamického návrhu, kdy se sledovaná veličina mění v závislosti na signálu).

Odhady střední hodnoty a rozptylu získáme z dat:

$$\hat{\mu} = \bar{y}_i = \frac{1}{M} \sum_{j=1}^M \bar{y}_{ij} = \frac{1}{M} \sum_{j=1}^M \left[\frac{1}{p} \sum_{k=1}^p y_{ijk} \right],$$

$$\hat{\sigma}^2 = S_{-i}^2 = \frac{1}{Mp} \sum_{j=1}^M \sum_{k=1}^p (y_{ijk} - \bar{y}_i)^2.$$

Přitom hodnoty y_{ijk} značí pozorování k -té replikace v i -tém bodě vnitřního pole při j -té kombinaci rušivých faktorů, \bar{y}_i je průměr vypočtený ze všech pozorování v i -tém bodě a S_i^2 značí výběrový rozptyl.

Chceme-li v procesu dosáhnout co nejmenší střední hodnoty (a zároveň co nejmenší variability), použijeme charakteristiku zvanou „menší je lepší“, danou vzorcem

$$z_i = -10 \log_{10} \frac{1}{Mp} \sum_{j=1}^M \sum_{k=1}^p y_{ijk}^2.$$

Toto lze použít v případě, kdy hodnoty responzní veličiny jsou kladné a cílová hodnota je nula. Naopak, chceme-li dosáhnout co největší střední hodnoty, použijeme kritérium „větší je lepší“, dané vztahem

$$z_i = -10 \log_{10} \frac{1}{Mp} \sum_{j=1}^M \sum_{k=1}^p \frac{1}{y_{ijk}^2}.$$

I v tomto případě musejí být hodnoty responzní veličiny kladné. Jakousi „oboustrannou varian-tou“, kdy požadujeme, aby střední hodnota byl rovna nějaké předem dané noiminální hodnotě, představuje kritérium založené na výrazu

$$z_i = -10 \log_{10} \frac{\bar{y}_i^2}{S_i^2}.$$

Při hledání robustního návrhu řešíme dvě úlohy. Nejdříve zkoumáme, které říditelné faktory mají vliv na charakteristiku variability (ať už jde o poměr signál/šum či logaritmus rozptylu) a určujeme úrovně vlivných faktorů, při nichž bude variabilita redukována. V dalším kroku zkoumáme vliv říditelných faktorů na průměr a očekáváme, že najdeme aspoň jeden faktor, který nemá zároveň vliv na variabilitu. Pomocí nastavení tohoto faktoru, případně dalších, upravíme střední hodnotu procesu. Vzhledem k tomu, že pro vnitřní pole se často využívá dílčí faktoriální návrh, dochází ke směřování některých efektů. Taguchi obchází tento problém tím, že předem předpokládá existenci pouze vybraných interakcí (nebo naopak žádné interakce nepředpokládá) a přiřazení sloupců matice návrhu jednotlivých faktorům či interakcím provádí tak, aby ke smíšení vybraných efektů nedocházelo. To znamená, že se vychází z určité předběžné znalosti procesu.

Taguchiho přístup se dá bez rizika použít tam, kde jsme si jisti, že interakce mezi faktory jsou zanedbatelné či dokonce neexistují. Analýza rozptylu je silně založena na předpokladu normality, což veličina s^2 nesplňuje. Přesto ale aplikovaná analýza rozptylu na s^2 může cokoliv objevit pro snižování úrovně variability podle úrovní řízených i neřízených faktorů. Aplikace Taguchiho návrhů je v praxi mezi inženýry hodně oblíbená, i když ze strany matematických statistiků je silně jeho přístup kritizován, že není zcela korektní a hlavně neumí vyhodnotit případné interakce mezi říditelnými a rušivými faktory, což lze v rámci faktoriálních návrhů např. řešit pomocí náhodných faktorů zapojených do experimentu.

Kapitola 6

DOE v MINITABu

6.1 Program MINITAB

Minitab je počítačový program určený pro statistickou analýzu dat. Byl vyvinut na Pensylvánské státní universitě (Pensylvania State University) týmem vědců v roce 1972. Program byl původně byl určen pro podporu výuky statistiky, nicméně je určen pro uživatele jak z oblasti praxe, tak z vysokých škol. Software nabízí řešení pro uživatele všech úrovní znalostí statistiky. Jeho výkon a jednoduchost používání z něho učinily vedoucí software na světě v oblasti výuky statistiky, analýzy dat a v neposlední řadě i v oblasti zlepšování kvality.

Současná verze Minitab 16 obsahuje prostředí pro práci s daty (ve formě tabulkového editoru), pro základní popisnou statistickou analýzu, pro tvorbu modelů, pro práci s časovými řadami, návrhy experimentů a mnoho dalších funkcí. Velkou podporu má grafické znázornění dat a výsledků analýz. Program umožňuje export dat a výsledků analýz do programů Microsoft Office (Excel, PowerPoint a Word). Kromě možnosti práce s vlastními daty nabízí i mnoho předpřipravených komentovaných příkladů a cvičení.

Zde je stručný přehled některých funkcí a vlastností, které v Minitabu naleznete:

Správa dat a souborů Aplikace Minitab přijímá data v mnoha formátech souborů a uchovává Vaši práci v jediném, automaticky uspořádaném souboru projektu.

Nejmodernější grafy s možností editace Jasné a přehledné grafy se snadno vytvářejí pro jakýkoliv účel. Úpravy se provádí intuitivně pomocí myši.

Popisná statistika Široká nabídka jednoduchých, ale výkonných nástrojů, které vám pomohou s prováděním rychlých vyhodnocení a srovnání.

Regresní analýza Lineární, polynomická, logistická a další regresní modely pomáhají nalézt a popsat vztahy mezi odezvou a jedním nebo více prediktory.

Analýza rozptylu Srozumitelná sada nástrojů obsahující nástroje ANOVA, GLM, analýzu středních hodnot a další.

Statistické řízení procesů Získejte kontrolu nad svými procesy pomocí sady regulačních diagramů a dalších nástrojů.

Nástroje řízení kvality Díky nástrojům pro analýzu způsobilosti procesů, Paretově grafu a dalším výkonným nástrojům budou procesy odpovídat Vaším očekáváním.

Analýza systému měření Analýza opakovatelnosti a reprodukovatelnosti měření (Gage R&R) a další související nástroje zajišťují kvalitu vašich dat, která analyzujete.

Plánované experimenty Široká nabídka dostupných experimentálních plánů, nyní včetně plánů D-optimal a plánů na bázi vzdálenosti.

Analýza spolehlivosti/analýza přežití Spočítejte si životnost Vámi používaných či prodávaných produktů, předpověďte četnost poruch, nastavte si parametry záruky atd.

Síla testu a rozsah vzorků Sbírejte správné množství dat pro danou situaci. Umožňuje použít Power křivky k vizualizaci vztahu mezi rozsahem vzorku a silou testu.

Multivarianční analýza Analýza hlavních komponent, analýza klastrů a další vícerozměrné nástroje.

Časové řady a předpovědi Rozsáhlá sbírka analýz souvisejících s časem, například analýza trendů a dekompozice. Nyní s auto, částečnou nebo cross korelací.

Neparametrická statistika Umožňuje provádění nezbytných analýz i s daty, která nemají normální rozdělení.

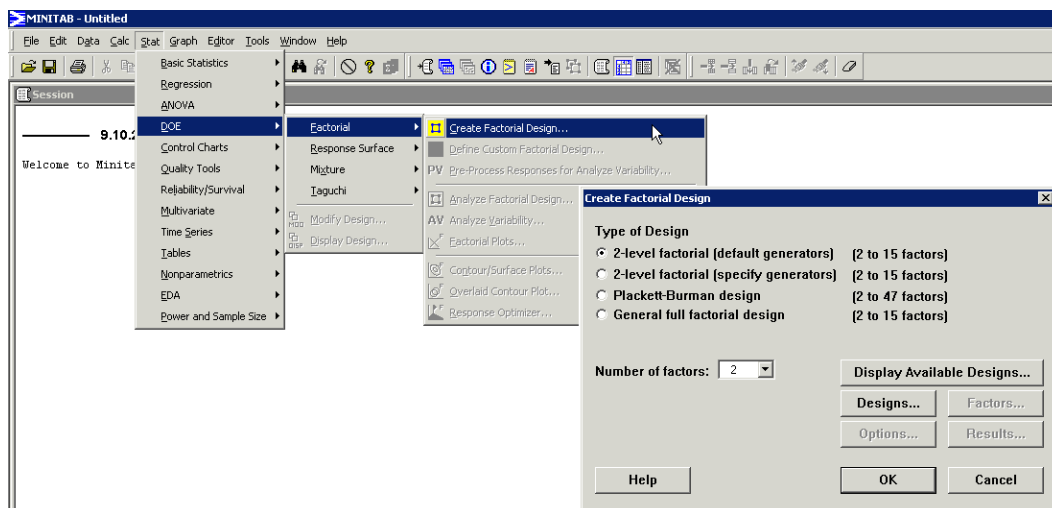
Tabulky Chi-kvadrát, Fisher a další nástroje na bázi tabulek.

Simulace a rozdělení Generujte náhodná čísla, vybírejte náhodné vzorky a mnohem víc.

Makra a přizpůsobitelnost Můžete automatizovat analýzy, které provádíte pravidelně, přizpůsobovat nabídky a nastavení nebo dokonce integrovat jiné programy pomocí objektů Minitab COM.

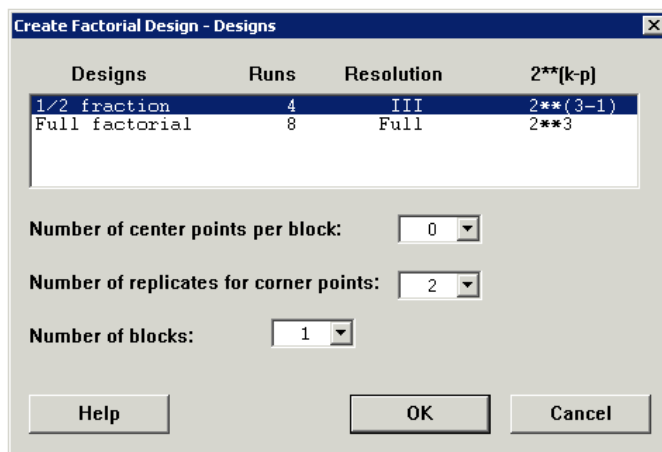
6.2 Vytvoření návrhu experimentu

Nástroje pro vytváření faktoriálních návrhů experimentů najdeme ve funkci *Stat* hlavní nabídky, v podnabídce *DOE > Factorial > Create Factorial Design*. V Obr. 6.6.2 jsou vidět i další možnosti (např. Taguchiho experimenty). Pro faktoriální návrhy lze nastavit počet faktorů a počet jejich úrovní.



Obr. 6.1: Vytvoření faktoriálního návrhu

Vytvoření dílčího experimentu můžeme provést stisknutím tlačítka *Designs...*. Po něm se nám objeví nabídkové okno, v němž lze vybrat kromě typu experimentu (úplný nebo dílčí) také počet replikací, počet centrálních bodů a případné bloky.

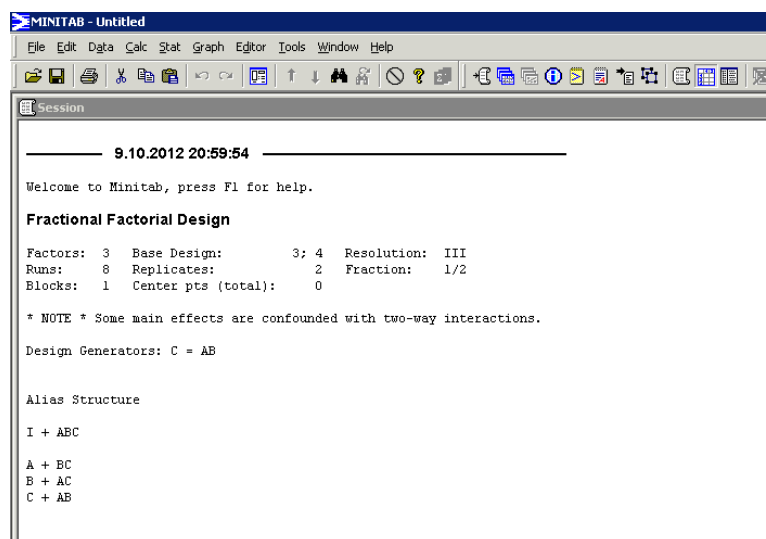


Obr. 6.2: Nastavení typu experimentu

Nastavení typu a vlastností navrhovaného experimentu má za následek předdefinování sloupců a řádků v tabulkovém editoru pro vstup dat.

6.3 Vstup dat

Po odsouhlasení výběru typu experimentu a jeho vlastností (stiskem tlačítka OK na panelu viz. Obr. 6.6.2) se ve výsledkovém okně objeví základní informace o navrhovaném experimentu:



Obrázek 6.1: Základní informace o experimentu

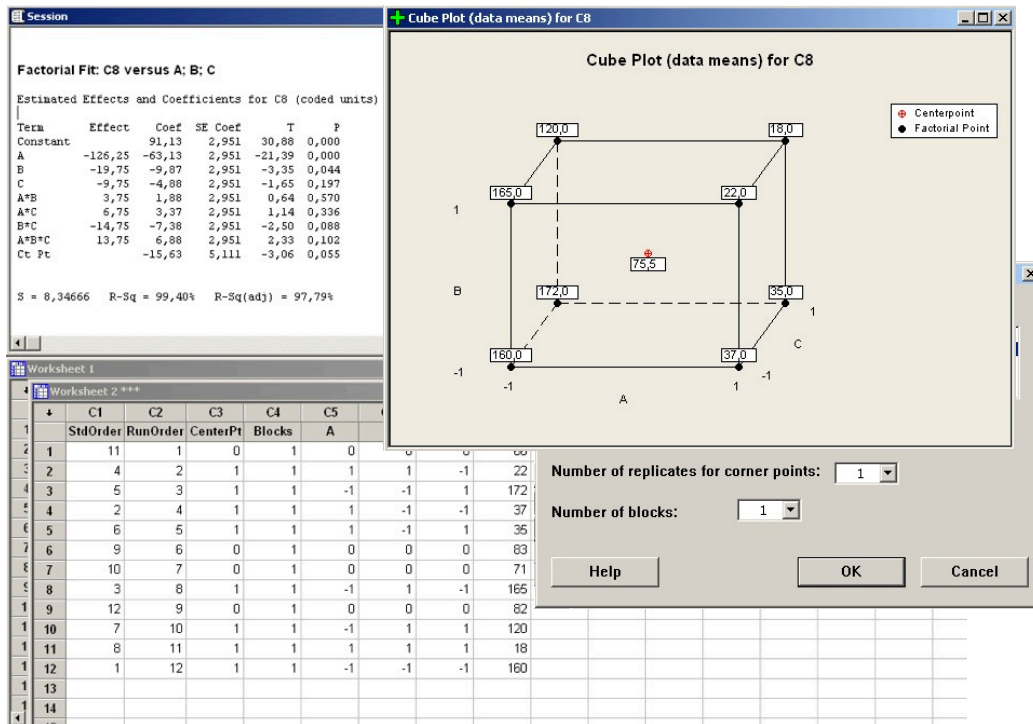
a v dolní části obrazovky najdeme předvyplněnou část řádků v tabulkovém editoru, sloužící pro vstup dat. Navržená měření obsahují všechny kombinace možných úrovní zúčastněných faktorů, přičemž každý řádek vlastního návrhu je zde zastoupen tolikrát, jaký je požadovaný počet replikací. Přitom kombinace úrovní faktorů jsou zde již znáhodněny, což je zřejmé z prvního sloupce původních pořadí (v uspořádané matici návrhu) a druhý sloupec udává, v jakém pořadí by se měla měření postupně uskutečňovat.

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	StdOrder	RunOrder	CenterPt	Blocks	A	B	C				
1	8	1	1	1	1	1	1				
2	5	2	1	1	-1	-1	1				
3	2	3	1	1	1	-1	-1				
4	1	4	1	1	-1	-1	1				
5	3	5	1	1	-1	1	-1				
6	4	6	1	1	1	1	1				
7	7	7	1	1	-1	1	-1				
8	6	8	1	1	1	-1	-1				
9											

Obrázek 6.2: Vstup dat

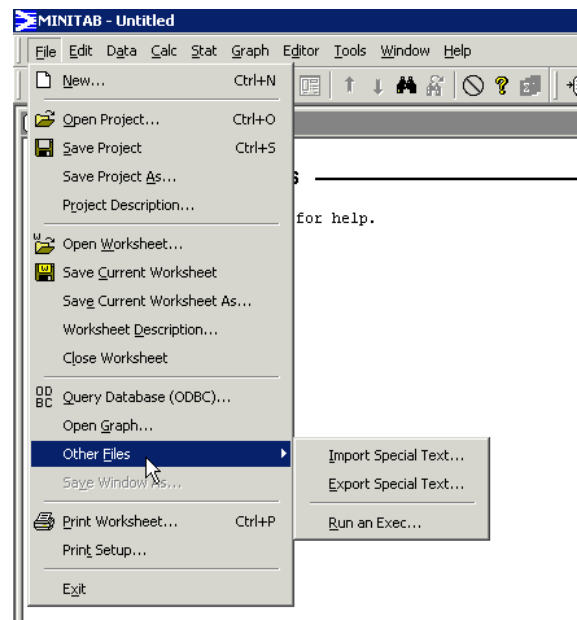
Naměřené hodnoty budeme zapisovat do sloupce C8, případně C9 a dalších.

Data lze do programu MINITAB také importovat ze souboru, vytvořeného například v MS



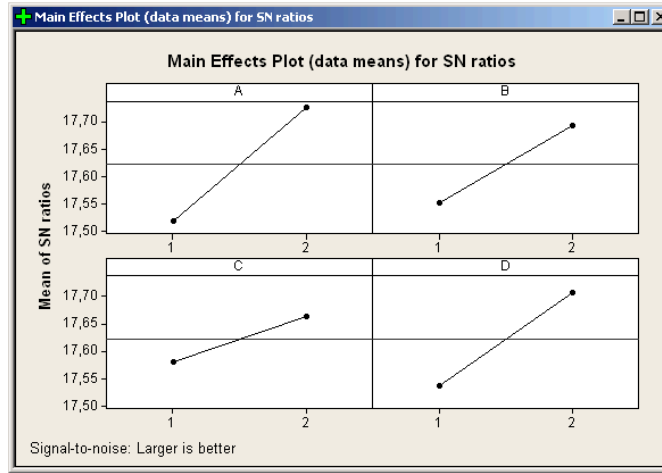
Obrázek 6.3: Vstup dat a jejich zobrazení ve formě krychle

Excelu nebo z databáze přes rozhraní ODBC. Vlastně lze importovat data z jakéhokoli programu, pokud budou v textovém formátu a oddělovačem ;.

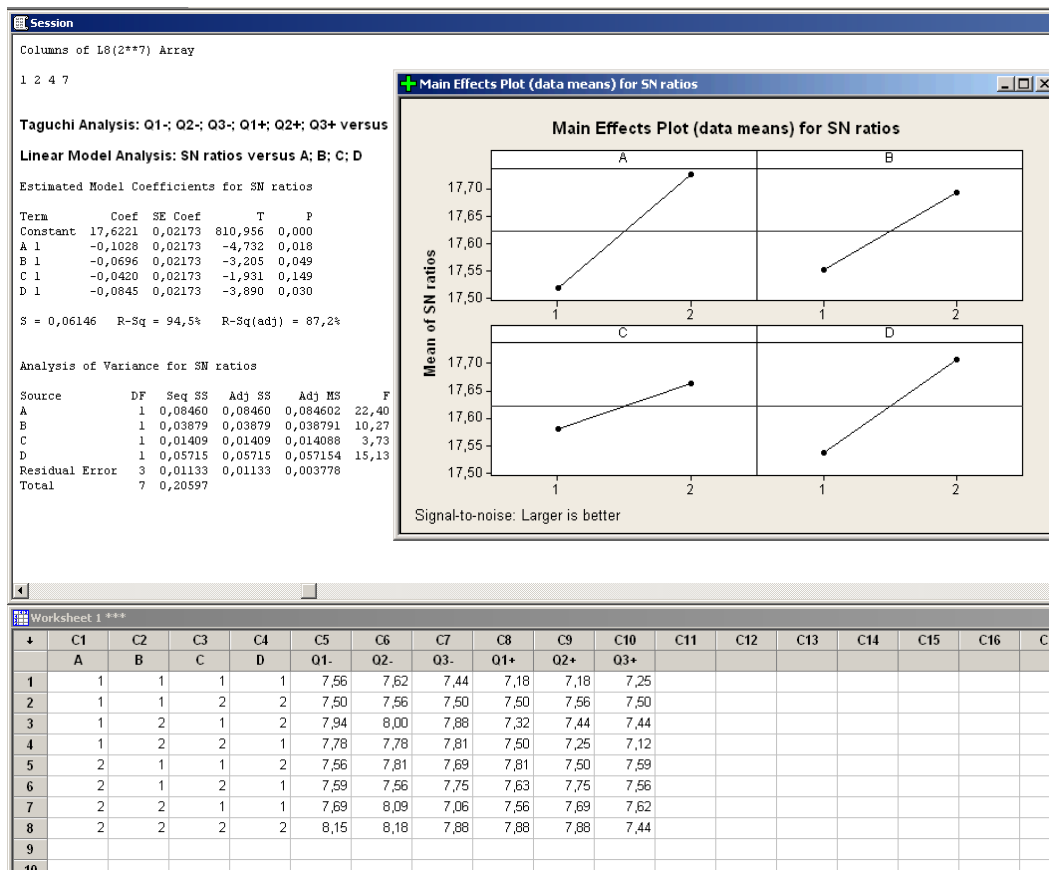


Obrázek 6.4: Vstup dat a jejich zobrazení ve formě krychle

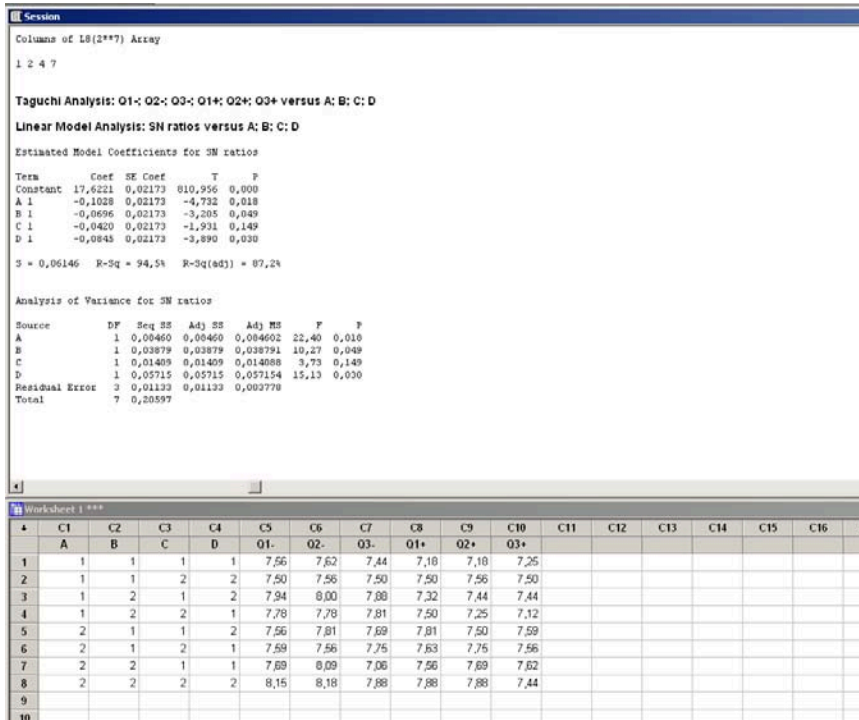
6.4 Vyhodnocení výsledku experimentu



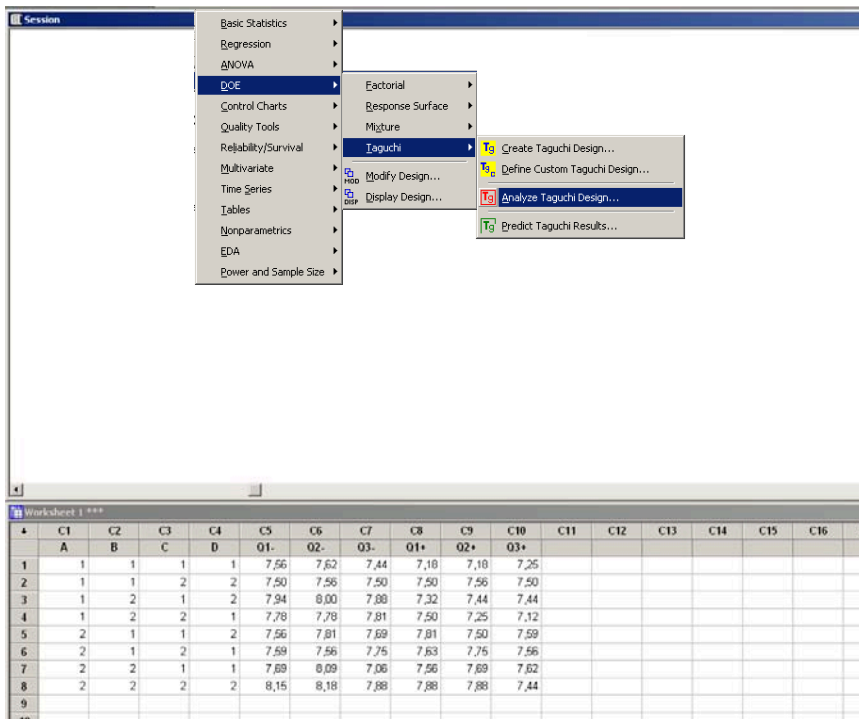
Obrázek 6.4: Hodnocení kontrolních diagramů



Obrázek 6.5: Hodnocení kontrolních diagramů



Obrázek 6.6: Hodnocení kontrolních diagramů



Obrázek 6.7: Hodnocení kontrolních diagramů

6.5 Řešený příklad z praxe

V této části si ukážeme konkrétní postup při návrhu a vyhodnocení experimentu pomocí programu Minitab. V dalším textu se omezíme pouze na činnosti, spojené s použitím tohoto programu. To znamená, že zde nebudeme popisovat všechny kroky návrhu experimentu, jako je například analýza procesu či detailní provedení zkoušek.

Účinnost a odolnost antibakteriálního účinku na upravených textiliích.

6.5.1 Návrh experimentu

Cílem experimentu je sledování úbytku účinnosti antibakteriálních úprav textilií v závislosti na počtu cyklů praní. Na základě analýzy procesu byly vytvořeny základní hypotézy, které by měl experiment potvrdit či vyvrátit:

- účinnost antibakteriální úpravy může být různá na různých antibakteriálních úpravách.
- účinnost se zřejmě mění s počtem pracích cyklů.

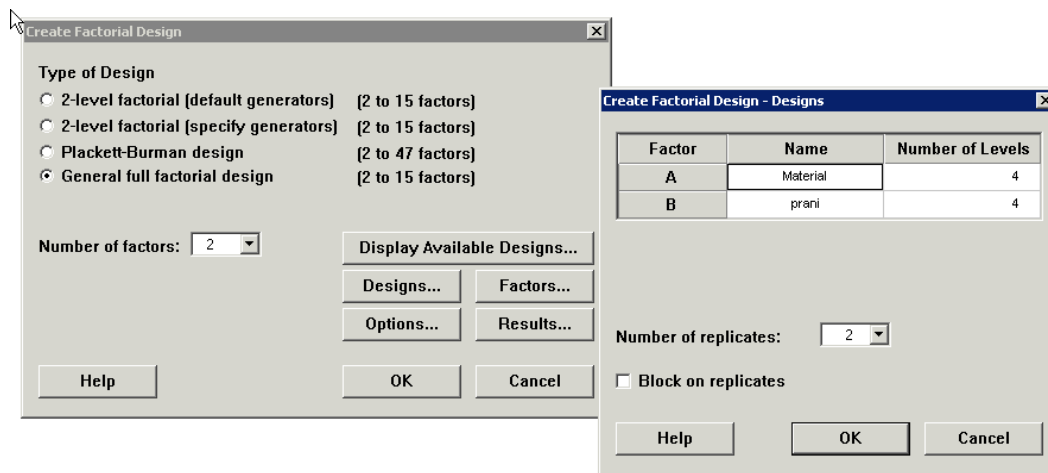
Zároveň byly při analýze vybrány čtyři druhy materiálu, na kterých bude experiment probíhat. Jako vhodný se ukázal návrh 2-faktorového experimentu, v němž

- první faktor je typ materiálu. Jsou uvažovány 4 typy materiálů, lišící se složením (poměrem bavlny a umělých vláken SeaCell, TreviraBioactive) nebo impregnací antibakteriálním přípravkem.
- druhým faktorem je počet cyklů praní. Měření bude prováděno před praním (0 cyklů), po 1 vyprání, po 25 cyklech a po 40 cyklech praní.
- od každého typu materiálu budou připraveny 2 vzorky pro každý ze čtyř uvažovaných počtů pracích cyklů. Tedy celkem budeme mít 32 vzorků.

Součástí návrhu je i způsob přípravy vzorků a provedení jednotlivých měření.

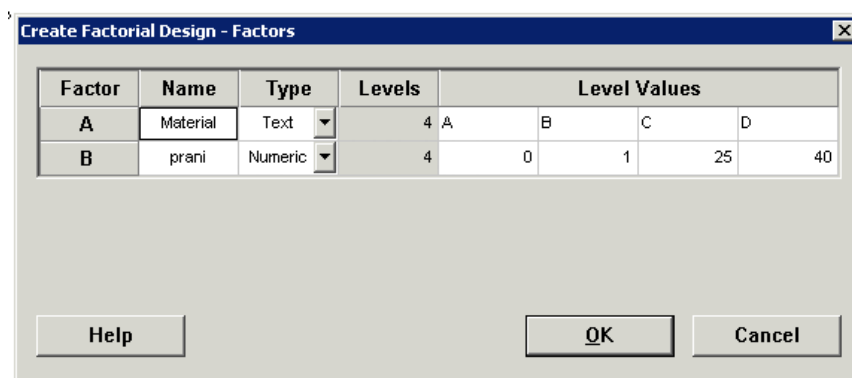
Náš vytvořený návrh experimentu nyní vložíme do programu Minitab, v němž bude probíhat i vyhodnocení. Spustíme program Minitab a budeme postupovat podle Obr. 6.6.2: v hlavní nabídce vybereme funkci statistiky (Stat), v ní metodu DOE, dále faktoriální návrh (Factorial) a jeho vytvoření (Create Factorial Design ...).

Nyní budeme pracovat s nabídkou v okně *Create Factorial Design*. Nejprve zaškrtneme možnost *General full fractional design*, neboť uvažujeme více než dvouúrovňový návrh. Po volbě tlačítka *Designs ...* se nám otevře okénko, v němž můžeme nastavit názvy faktorů (zvolili jsme *Material* a *prani*) a počty jejich úrovní (v našem případě jsou 4 pro oba faktory). Počet replikací (*Number of replicates*) nastavíme na hodnotu 2.



Obrázek 6.8: Vytvoření faktoriálního návrhu

Volba *Factors ...* nám dovolí označit hodnoty jednotlivých úrovní faktorů. Zde jsme označili jednotlivé druhy materiálů velkými písmeny: A (60% Bavlna, 20% SeaCell active, 20% SeaCell pure), B (50% Bavlna, 50% TreviraBioactive), C (50% Bavlna, 30% Smartcell sensitive, SeaCell pure) a D (100% Bavlna impregnovaná antibakteriálním prostředkem SILVERPLUS® protection), počty cyklů praní potom hodnotami 0, 1, 25 a 40.



Obrázek 6.9: Nastavení faktorů

Dále zvolíme randomizaci návrhu. To proto, abychom vyloučili možnou závislost měření danou pořadím vzorků nebo počty cyklů praní. Randomizaci nastavíme zaškrtnutím políčka v okně

Options Ještě použijeme volbu *Results* ..., v níž nastavíme možnosti výstupu výsledků na *Summary table and design table*. V tomto formátu se nám výsledky ukáží na ploše *sessions* a bude možné je tak i vytisknout či uložit. Nyní ukončíme návrh stiskem tlačítka *OK*. Vzápětí se nám ukáže tabulka návrhu jak v poli *sessions*, tak i v tzv. *Worksheetu*. Do něho budeme zapisovat výsledky měření, postupně řádek po řádku. Známé pořadí se zde projevuje tím, že jednotlivé řádky ve worksheetu jsou už označeny úrovněmi faktorů (ve sloupcích *Matreriál* a *prani* v náhodném pořadí).

The image shows two windows from the MINITAB software interface. The 'Session' window on the left displays a list of 32 runs with columns 'Run', 'Blk', 'A', and 'B'. The 'Worksheet 5 ***' window on the right displays a table with columns 'StdOrder', 'RunOrder', 'PtType', 'Blocks', 'Material', 'prani', 'C7', and 'C8'. The data in the worksheet is as follows:

	C1	C2	C3	C4	C5-T	C6	C7	C8
	StdOrder	RunOrder	PtType	Blocks	Material	prani		
1	6	1	1	1	B	1		
2	30	2	1	1	D	1		
3	25	3	1	1	C	0		
4	1	4	1	1	A	0		
5	31	5	1	1	D	25		
6	29	6	1	1	D	0		
7	26	7	1	1	C	1		
8	2	8	1	1	A	1		
9	15	9	1	1	D	25		
10	14	10	1	1	D	1		
11	24	11	1	1	B	40		
12	4	12	1	1	A	40		
13	16	13	1	1	D	40		
14	7	14	1	1	B	25		
15	21	15	1	1	B	0		
16	10	16	1	1	C	1		
17	19	17	1	1	A	25		
18	18	18	1	1	A	1		
19	32	19	1	1	D	40		
20	27	20	1	1	C	25		
21	20	21	1	1	A	40		
22	13	22	1	1	D	0		
23	17	23	1	1	A	0		
24	5	24	1	1	B	0		
25	12	25	1	1	C	40		
26	3	26	1	1	A	25		
27	11	27	1	1	C	25		
28	28	28	1	1	C	40		
29	22	29	1	1	B	1		
30	9	30	1	1	C	0		
31	8	31	1	1	B	40		
32	23	32	1	1	B	25		
33								
34								

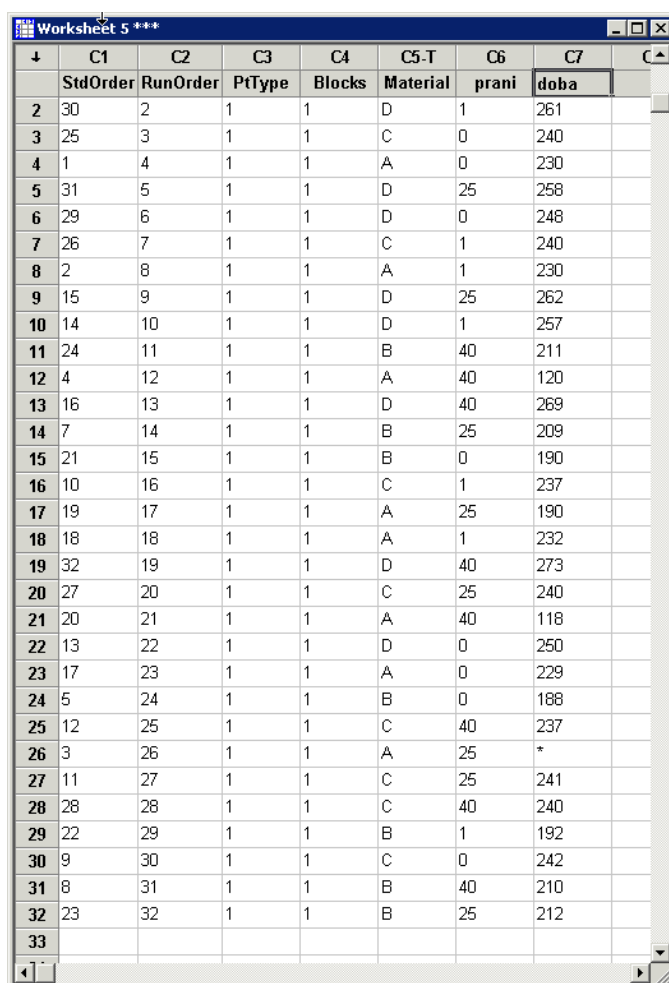
Obrázek 6.10: Worksheet

6.5.2 Příprava vzorků a měření

Vzorky, které budou procházet pracím cyklem, budou vyprány podle doporučeného pracího postupu v souladu s normou ČSN EN ISO 6330 (80 0821).

Po praní budou testovány změny antibakteriálních vlastností antibakteriálně upravených textilií. Připravené a vyprané vzorky budou vystaveny vlivu kvasného roztoku a bude sledován proces fermentace. Během procesu fermentace kvasinky přeměňují jednoduché cukry (glukosa, fruktosa) na ethanol (alkohol) a CO₂ (oxid uhličitý), který uchází do okolního vzduchu. Princip experimentu je založen na sledování doby fermentace, po kterou se uvolňuje oxid uhličitý a odpovídá aktivitě kvasinek.

Do této tabulky zapíšeme naměřené hodnoty do sloupce C7, který navíc označíme jako *doba*.

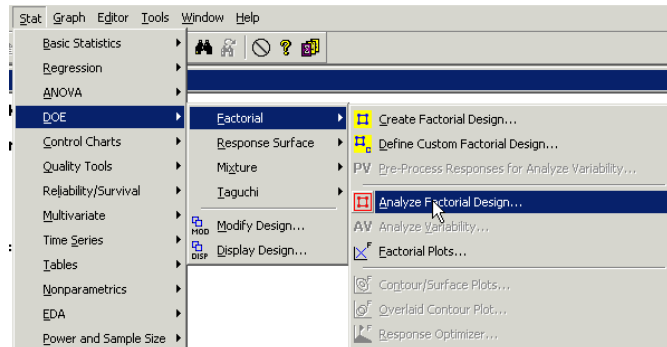


	C1	C2	C3	C4	C5-T	C6	C7
	StdOrder	RunOrder	PtType	Blocks	Material	prani	doba
2	30	2	1	1	D	1	261
3	25	3	1	1	C	0	240
4	1	4	1	1	A	0	230
5	31	5	1	1	D	25	258
6	29	6	1	1	D	0	248
7	26	7	1	1	C	1	240
8	2	8	1	1	A	1	230
9	15	9	1	1	D	25	262
10	14	10	1	1	D	1	257
11	24	11	1	1	B	40	211
12	4	12	1	1	A	40	120
13	16	13	1	1	D	40	269
14	7	14	1	1	B	25	209
15	21	15	1	1	B	0	190
16	10	16	1	1	C	1	237
17	19	17	1	1	A	25	190
18	18	18	1	1	A	1	232
19	32	19	1	1	D	40	273
20	27	20	1	1	C	25	240
21	20	21	1	1	A	40	118
22	13	22	1	1	D	0	250
23	17	23	1	1	A	0	229
24	5	24	1	1	B	0	188
25	12	25	1	1	C	40	237
26	3	26	1	1	A	25	*
27	11	27	1	1	C	25	241
28	28	28	1	1	C	40	240
29	22	29	1	1	B	1	192
30	9	30	1	1	C	0	242
31	8	31	1	1	B	40	210
32	23	32	1	1	B	25	212
33							

Obrázek 6.11: Naměřené hodnoty doby fermentace

6.5.3 Analýza výsledků

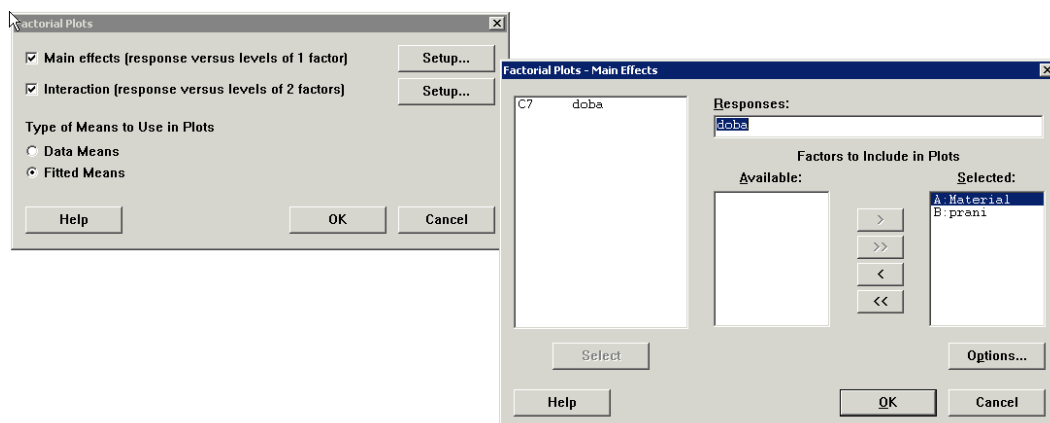
Nyní přikročíme k analýze výsledků experimentu. V hlavní nabídce použijeme opět cestu *Stat* → *DOE* → *Factorial*. Zde se nám objevily další dvě přístupné funkce: *Analyze Factorial Design ...* a *Factorial Plots ...*.



Obrázek 6.12: Vyhodnocení experimentu

Podívejme se nejprve na grafy efektů hlavních faktorů. Vybereme funkci *Factorial Plots ...*. Zde máme možnost zaškrtnout jednu nebo obě možnosti:

- graf hlavních efektů (*Main effects*) zobrazující střední hodnoty odezvy při jednotlivých úrovních hlavních faktorů. Můžeme si navíc vybrat, zda chceme střední hodnoty přímo z naměřených dat nebo odhadnuté střední hodnoty z modelu.
- graf interakcí (*Interaction*) zobrazující hodnoty středních hodnot odezvy pro různé úrovně druhého faktoru ve čtyřech křivkách, každé pro jednu hodnotu prvního faktoru.

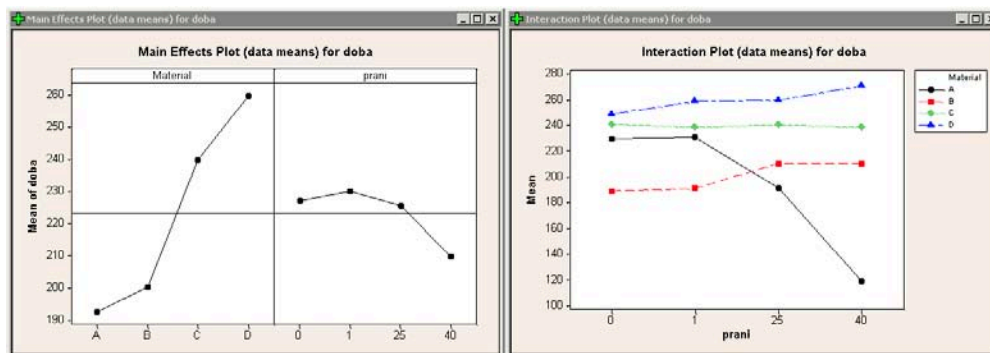


Obrázek 6.13: Grafy hlavních efektů a interakcí

Dále je třeba nastavit odezvu a faktory. K tomu slouží funkce *Setup ...*. Zde výběrem ze

seznamu proměnných zvolíme, která bude odezvou (doba), které hlavními faktory (Mereni, prani) a pro které kombinace faktorů chceme zobrazit interakce.

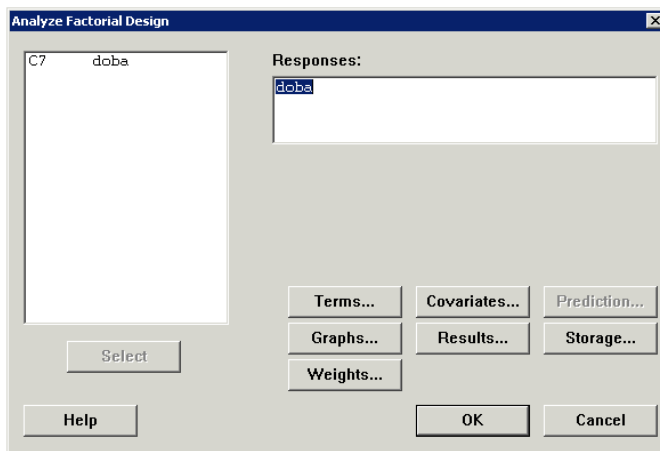
Po stisku tlačítka OK dostaneme zaškrtnuté grafy, z nichž lze vyčíst vlivy hlavních faktorů, případně přítomnost vlivu interakcí. Tyto grafy neposkytují kvantitativní údaje a tudíž na jejich základě lze pouze „získat dojem“ o vlivu jednotlivých faktorů či interakcí. Nicméně, jako vodítko pro další analýzu jsou užitečné.



Obrázek 6.14: Hlavní efekty faktorů a interakce

Z grafu nalevo je zřejmé, že vliv materiálu je poměrně velký, přičemž nejlepší výsledky vykazuje materiál D a nejhorsí materiál A. V případě materiálu A odolnost výrazně klesá s počtem pracích cyklů na rozdíl od ostatních tří materiálů, jak lze vidět na Obr. 6.13 napravo. Tento výrazný trend ovlivňuje zřejmě i graf hlavních efektů v pravé části levého grafu.

Z předchozích grafů je zřejmé, že hlavní faktory i jejich interakce mají vliv na odezvu. Jak velký a významný, to nám poví kvantitativní analýza. Tu vyvoláme volbou funkce *Analyze Factorial Design*



Obrázek 6.15: Analyze Factorial Design ...

Panel pro analýzu výsledků experimentu nabízí několik funkcí:

- *Terms ...* pro nastavení stupně interakce. V našem případě lze volit 1 (bez interakce) nebo 2 (interakce mezi dvěma uvažovanými faktory)
- *Covariates ...* je panel pro výběr kovariát v modelu. Tyto kovariáty musejí být uloženy jako sloupec v pracovním sešitě.
- *Graphs ...* umožňuje nastavení výstupních grafů výběrem z několika možností
- *Results ...* nastavuje tvar výstupu. Pokud nezvolíme nic (*Do not display*), je základem tabulka analýzy rozptylu (*ANOVA Table*) a k ní lze vybrat další informace (*covariate, coefficients, unusual observations*)
- *Storage ...* pro uložení reziduí, odhadů, parametrů modelu a dalších informací. Tyto údaje budou uloženy jako sloupce tabulky v pracovním sešitu (*Worksheet*).
- *Weights ...* je panel pro výběr vah pro vážený odhad odezvy. Tyto váhy musejí být uloženy jako sloupec v pracovním sešitě.

V našem případě pouze nastavíme typ výstupu s rezidui a stiskneme OK (odezva a faktory spolu s interakcemi už máme nastaveny z předchozího kroku (nastavení pro grafy hlavních efektů a interakcí). Poté dostaneme přehled výsledků kvantitativní analýzy.

General Linear Model: doba versus Material; prani

Factor	Type	Levels	Values
Material	fixed	4	A; B; C; D
prani	fixed	4	0; 1; 25; 40

Náš model uvažujeme ve tvaru

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

kde y_{ij} je doba fermentace pro i -tou úroveň faktoru materiál a j -tou úroveň počtu cyklů praní, μ je společná úroveň pro oba faktory, α_i označuje vliv i -té úrovně faktoru *material*, β_j označuje vliv j -té úrovně faktoru *prani*, $(\alpha\beta)_{ij}$ je vliv interakce i -té a j -té úrovně faktorů a ϵ_{ij} představuje náhodnou chybu při i -té a j -té úrovni odpovídajících faktorů (reziduum). Statistickou významnost vlivu faktorů a jejich interakce vyhodnotíme pomocí analýzy rozptylu:

Analysis of Variance for doba, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Material	3	24804,0	24804,0	8268,0	4724,57	0,000
prani	3	1983,3	1983,3	661,1	377,76	0,000
Material*prani	9	15232,2	15232,2	1692,5	967,13	0,000
Error	16	28,0	28,0	1,8		
Total	31	42047,5				

S = 1,32288 R-Sq = 99,93% R-Sq(adj) = 99,87%

Z tabulky analýzy rozptylu je vidět, že vliv obou faktorů i jejich interakce jsou statisticky významné. O tom svědčí prakticky nulové p -hodnoty v posledním sloupci. Hodnoty koeficientu determinace $R-Sq$ jsou blízké jedné, což vypovídá o tom, že náš model vysvětluje téměř všechnu variabilitu naměřených dat a tedy je dobrým modelem. Odhady jednotlivých efektů (koeficientů modelu) jsou v další části textového výstupu:

Term		Coef	SE Coef	T	P
Constant		223,125	0,234	954,12	0,000
Material					
A		-30,5000	0,4050	-75,30	0,000
B		-23,0000	0,4050	-56,78	0,000
C		16,5000	0,4050	40,74	0,000
prani					
0		4,2500	0,4050	10,49	0,000
1		7,1250	0,4050	17,59	0,000
25		1,8750	0,4050	4,63	0,000
Material*prani					
A	0	32,6250	0,7016	46,50	0,000
A	1	30,7500	0,7016	43,83	0,000
A	25	-4,5000	0,7016	-6,41	0,000
B	0	-14,3750	0,7016	-20,49	0,000
B	1	-16,7500	0,7016	-23,88	0,000
B	25	7,5000	0,7016	10,69	0,000
C	0	-3,8750	0,7016	-5,52	0,000

C	1	-6,7500	0,7016	-9,62	0,000
C	25	-1,0000	0,7016	-1,43	0,173

Zde *Constant* odpovídá společnému efektu μ , *material* A, B, C odpovídá efektům $\alpha_1, \alpha_2, \alpha_3$, *prani* 0, 1, 25 odpovídá efektům $\beta_1, \beta_2, \beta_3$ a ve spodní polovině tabulky jsou hodnoty efektů interakcí. Všimněme si, že výsledková tabulka neobsahuje hodnoty efektů pro materiál D a počet cyklů praní 40. Tyto efekty lze snadno dopočítat z podmínky $\sum_{i=1}^4 \alpha_i = 0$ a $\sum_{j=1}^4 \beta_j = 0$. Odtud dostáváme $\alpha_4 = 37,000$ a $\beta_4 = -13,250$.

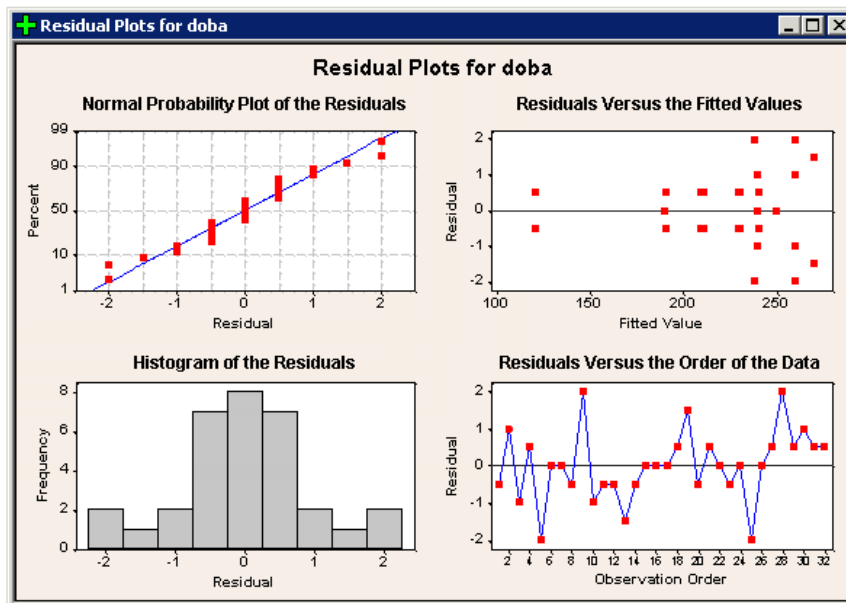
Následující odhady středních hodnot odezvy pro jednotlivé úrovně faktorů a pro interakce jsme už viděli graficky znázorněné v Obr. 6.14:

Least Squares Means for doba

Material		Mean	SE Mean
A		192,6	0,4677
B		200,1	0,4677
C		239,6	0,4677
D		260,1	0,4677
prani			
	0	227,4	0,4677
	1	230,3	0,4677
	25	225,0	0,4677
	40	209,9	0,4677
Material*prani			
A	0	229,5	0,9354
A	1	230,5	0,9354
A	25	190,0	0,9354
A	40	120,5	0,9354
B	0	190,0	0,9354
B	1	190,5	0,9354
B	25	209,5	0,9354
B	40	210,5	0,9354
C	0	240,0	0,9354
C	1	240,0	0,9354
C	25	240,5	0,9354

C	40	238,0	0,9354
D	0	250,0	0,9354
D	1	260,0	0,9354
D	25	260,0	0,9354
D	40	270,5	0,9354

Volba *Graphs ...* nám nabízí řadu grafů zobrazujících především rezidua. Zvolíme „čtyři v jednom“ (*Four in one*) obsahující: normální pravděpodobnostní papír pro rezidua (*Normal Probability Plot of the Residuals*), graf reziduí pro odhadnuté hodnoty (*Residuals Versus the Fitted Values*), histogram reziduí (*Histogram of the Residuals*) a graf reziduí pro jednotlivá měření (*Residuals Versus the Order of the Data*). Výstup je na Obr. 6.16.



Obrázek 6.16: Analýza reziduí

Z histogramu vidíme, že v krajních hodnotách zjištěných reziduí (v bodech -2 a +2) existuje více pozorování, než by odpovídalo normálnímu rozdělení. V pravém dolním grafu na obr. 6.16 je vidět, že existují čtyři měření v nichž rezidua v našem modelu nabývají extrémních hodnot. Tato měření jsou identifikována v textovém výstupu jako *neobvyklá pozorování* (*unusual observations*):

Unusual Observations for doba

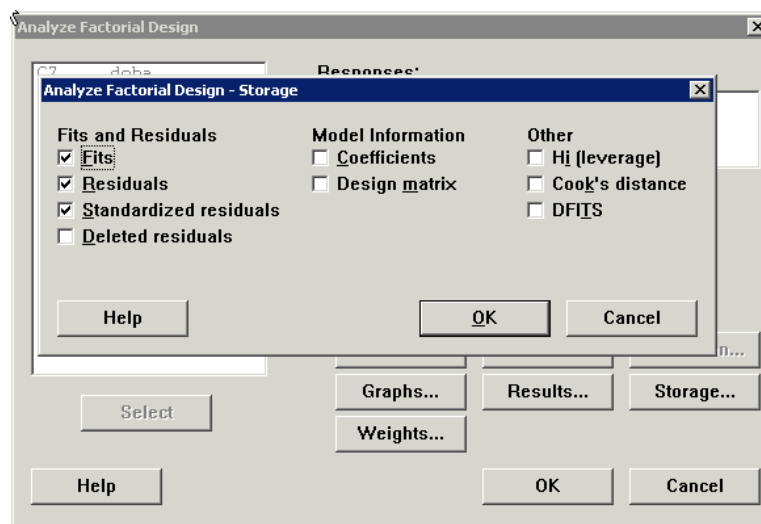
Obs	doba	Fit	SE Fit	Residual	St Resid
5	258,000	260,000	0,935	-2,000	-2,14 R
9	262,000	260,000	0,935	2,000	2,14 R

25	236,000	238,000	0,935	-2,000	-2,14 R
28	240,000	238,000	0,935	2,000	2,14 R

R denotes an observation with a large standardized residual.

Grafy na Obr. 6.16 nám ukazují, že by mohl být problém s normalitou reziduí. Provedeme tedy ještě test normality reziduí.

Uložíme nejprve rezidua do datové tabulky nastavením funkce *Storage ...* v panelu analýzy faktoriálního experimentu:

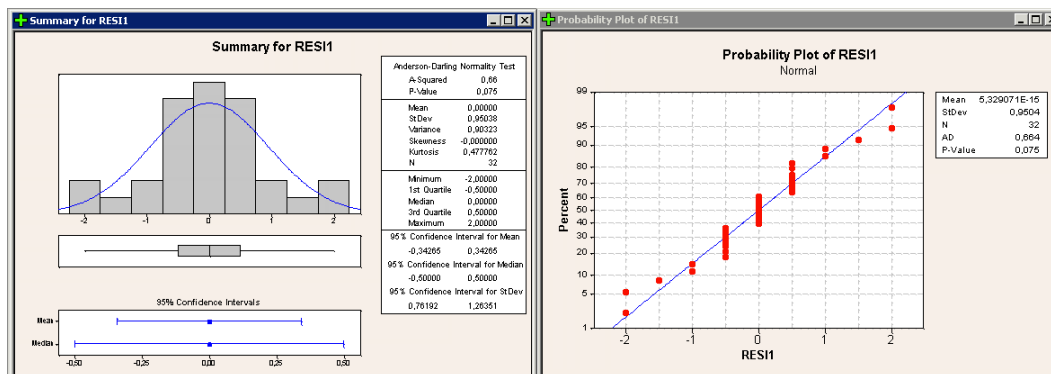


Obrázek 6.17: Uložení reziduí a predikce

Vybrali jsme uložení odhadovaných hodnot (*Fits*), reziduí (*Residuals*) a standardizovaných reziduí (*Standardized Residuals*). Spočtená data se uloží do sloupců tabulky pod názvy FITS1, RESI1 a SRES1.

Test normality nalezneme v hlavní nabídce pod položkou *Stat* → *Basic statistics*. Zde můžeme vybrat buď funkci *Normality test ...* nebo *Graphical Summary ... Normality test ...* nabízí několik testů normality: Anderson-Darlingův test, test Ryan-Joiner (obdoba Shapiro-Wilkova testu) a Kolmogorov-Smirnovův test. V našem případě, kdy neznáme střední hodnotu ani rozptyl, je nejvhodnější použít Anderson-Darlingův test. Jeho výsledek je zobrazen na Obr. 6.18 vpravo. Funkce *Graphical Summary ...* nám poskytne podobný výsledek pouze s tím rozdílem, že data zobrazuje v histogramu a krabicovém grafu (viz Obr. 6.18 vlevo).

Z výsledků je zřejmé, že hypotézu o normalitě nelze zamítnout na hladině významnosti 7,5%.



Obrázek 6.18: Test normality

6.5.4 Závěr

Cílem experimentu bylo sledování úbytku účinnosti antibakteriálních úprav textilií v závislosti na počtu cyklů praní. Na základě výsledku měření lze konstatovat, že účinnost antibakteriálních úprav se výrazně liší v závislosti na použitých materiálech. Pokles účinnosti v závislosti na počtu pracích cyklů je pouze „optická“ (viz Obr. 6.14 vlevo). Zde se projevuje silný vliv interakce *materiál*praní*. Pokles účinnosti s počtem pracích cyklů je výrazný pouze v případě materiálu A (60% Bavlna, 20% SeaCell active, 20% SeaCell pure). U ostatních použitých materiálů je tomu spíše naopak (viz Obr. 6.14 vpravo).

Literatura

- [1] Jarošová, E.: Navrhování experimentů, VŠE , 1998
- [2] Jarošová, E.: Navrhování experimentů a jejich analýza, Česká společnost pro jakost, 2007
- [3] Likeš, J.: Navrhování průmyslových experimentů, SNTL, 1968
- [4] Michálek, J., Křepela, J.: DOE – Navrhování a vyhodnocování statistických experimentů, ČSJ, 2009
- [5] Montgomery, D. C.: Design and Analysis of Experiments , 5th ed., J.Wiley, 2002
- [6] Roy, R.K.: Design of experiments using the Taguchi aproach. J.Wiley&Sons, New York, 2001
- [7] Ryan, P. R.: Modern Experiment Design, J. Wiley, 2007
- [8] Tošenovský, J.: Plánování experimentů, VŠB-TU Ostrava, 2012
- [9] ČSN ISO 3534-3:2001 Statistika – Slovník a značky – Část 3: Navrhování experimentů (01 0216)