

POUŽÍVATELSKÁ PRÍRUČKA K JMP

Štatistika a návrh experimentov (DOE)

Pre JMP Student Edition 19.x

Príručka je navrhnutá tak, aby fungovala dvoma spôsobmi: ako stručný výklad štatistických a experimentálnych princípov a zároveň ako návod „klik za klikom“. Každá metóda obsahuje cieľ, vhodné použitie, cestu v menu JMP, odporúčaný pracovný postup, interpretáciu výstupu a upozornenia na časté chyby.

Prvok	Význam
V JMP:	Presná alebo odporúčaná cesta v menu. Názvy položiek ostávajú v angličtine.
Tip	Efektívnejší postup alebo odporúčanie pre výučbu.
Cvičenie	Úloha viazaná na priložený dátový súbor.
Kontrola výsledku	Orientačné hodnoty na overenie práce; drobné rozdiely môžu vzniknúť zaokrúhlením.

Rozsah a aktuálnosť

Postupy sú zostavené pre aktuálnu vetvu JMP Student Edition 19.x. Menšie rozdiely sa môžu objaviť medzi systémami Windows a macOS alebo po aktualizácii opráv. Student Edition je určená študentom, vyučujúcim a akademickým výskumníkom; prístup sa pravidelne obnovuje po overení akademickej príslušnosti. Poskytuje plnohodnotné prostredie vhodné od základnej štatistiky po pokročilé modelovanie a DOE.

Dôležité: Ak sa názov položky v menu mierne líši, použite vyhľadávanie v Help alebo príkaz Help > Search. Pri DOE sa vždy riadte skriptom Model uloženým v tabuľke návrhu, pretože zachováva správnu špecifikáciu modelu.

Výsledky vzdelávania s použitím JMP

- správne založiť alebo importovať dátovú tabuľku a priradiť premenným dátové a modelovacie typy;
- vykonať a interpretovať deskriptívnu analýzu, grafickú exploráciu a základné testy hypotéz;
- použiť t-testy, ANOVA, koreláciu, lineárnu a logistickú regresiu vrátane diagnostiky;
- navrhnuť úplný a frakčný faktoriálny experiment, screeningový experiment, vlastný optimálny návrh a plán odozvovej plochy;
- analyzovať hlavné efekty a interakcie, pracovať s Prediction Profiler a formulovať vecný záver s obmedzeniami.

Obsah

Označenie	Kapitola
ČASŤ I	Prostredie JMP a práca s dátami
1	Orientácia v JMP
2	Dátová tabuľka a príprava dát
3	Exploračná analýza a vizualizácia
ČASŤ II	Štatistická inferencia
4	Logika testovania a intervalov spoľahlivosti
5	Porovnávanie priemerov
6	ANOVA a viacfaktorové modely
7	Korelácia a regresia
8	Kategorické dáta a logistická regresia
ČASŤ III	Návrh experimentov (DOE)
9	Princípy DOE a pracovný postup
10	Úplný faktoriálny plán
11	Frakčný faktoriálny a screeningový plán
12	Custom Design
13	Plocha odozvy a optimalizácia
14	Analýza, diagnostika a reporting DOE
ČASŤ IV	Cvičenia a pomôcky
15	Súhrnné cvičenia s riešeniami
16	Najčastejšie chyby
17	Rýchla mapa menu
18	Slovník pojmov a zdroje

ČASŤ I

Prostredie JMP a práca s dátami

1. Orientácia v JMP

JMP je interaktívne štatistické prostredie. Výsledok analýzy nie je statická tabuľka: výber riadkov v dátach sa zvýrazní v grafoch a naopak. Táto väzba pomáha odhaľovať skupiny, odľahlé pozorovania a štruktúru dát.

1.1 Základné objekty

Objekt	Úloha	Praktická poznámka
Data Table	Ukladá riadky, stĺpce, vlastnosti premenných a skripty analýz.	Jeden riadok má spravidla predstavovať jednu experimentálnu alebo pozorovaciu jednotku.
Platform launch window	Dialóg na priradenie premenných k analytickým rolám Y, X, By, Weight a pod.	Pred spustením skontrolujte modelovací typ ikonou pri názve stĺpca.
Report window	Interaktívny výstup s grafmi, tabuľkami a ponukou červeného trojuholníka.	Kliknutím na červený trojuholník pridávate testy, diagnostiku a ukladáte výsledky.
Journal / Project	Spája výstupy, odkazy a komentáre do jedného pracovného celku.	Vhodné na odovzdanie seminárnej práce alebo výučbového balíka.
Table Scripts	Uložené príkazy na zopakovanie analýzy.	Pri DOE sú skripty Design a Model zásadné; nemažte ich.

1.2 Červený trojuholník

Červený trojuholník je kontextová ponuka JMP. Nachádza sa pri názve celej analýzy aj pri jednotlivých častiach výstupu. Obsahuje doplnkové grafy, testy, ukladanie predikcií, skripty a možnosti zobrazenia. Základný výstup často neobsahuje všetko, čo je potrebné na diagnostiku.

Pravidlo práce: Po spustení každej analýzy urobte tri kroky: (1) prezrite graf, (2) otvorte červený trojuholník a zapnite potrebnú diagnostiku, (3) formulujte vecný záver, nie iba „ $p < 0,05$ “.

1.3 Pomoc, ukážkové dáta a vyhľadávanie

V JMP: Help > Sample Data Folder

Priečink ukážkových dát obsahuje tabuľky pre väčšinu platforiem. Pre DOE hľadajte podpriečinok Design Experiment. Dokumentáciu otvoríte cez Help a online vyhľadávanie; krátke návody sú dostupné aj v JMP Learning Library.

1.4 Odporúčané nastavenie výučby

- Používajte anglické názvy ponúk aj v slovenskom texte; umožní to ich ľahšie nájdenie v dokumentácii a videách.
- Zobrazujte Column Info pred každou analýzou, aby sa upevnilo rozlíšenie Data Type a Modeling Type.
- Ukladajte postup ako table script: červený trojuholník > Save Script > To Data Table.
- Pred odovzdaním výsledkov uložte dátovú tabuľku vo formáte .jmp, nie iba exportovaný obrázok.

2. Dátová tabuľka a príprava dát

2.1 Riadky, stĺpce a jednotka pozorovania

Dobrá dátová tabuľka má jednoznačnú jednotku pozorovania. Pri klasickom usporiadaní je každý riadok jedna osoba, výrobok, vzorka alebo experimentálny beh a každý stĺpec jedna premenná. Opakované merania môžu byť v širokom tvare (viac stĺpcov pre časy) alebo v dlhom tvare (stĺpec Čas a stĺpec Hodnota); výber závisí od analýzy.

Kontrolná otázka: Dokážete jednou vetou dokončiť tvrdenie „Jeden riadok v tejto tabuľke predstavuje ...“? Ak nie, tabuľka pravdepodobne nemá správnu štruktúru.

2.2 Data Type a Modeling Type

V JMP: Cols > Column Info

Nastavenie	Možnosti	Použitie
Data Type	Numeric, Character, Row State	Určuje fyzický spôsob uloženia. Dátum a čas sú numerické údaje so špeciálnym formátom.
Modeling Type	Continuous, Ordinal, Nominal, Multiple Response, Unstructured Text, Vector	Určuje, ako JMP premennú použije v analýze.
Format	Počet desatinných miest, dátum, percentá a pod.	Mení zobrazenie, nie samotnú uloženú hodnotu.
Column Properties	Formula, Value Labels, Units, Spec Limits, Notes, Design Role...	Pridáva metadáta a správanie stĺpca.

Premenná	Data Type	Modeling Type	Prečo
Výška v cm	Numeric	Continuous	Rozdiely a priemery majú zmysel.
Ročník 1–5	Numeric alebo Character	Ordinal	Poradie má zmysel, vzdialenosť nemusí byť rovnaká.
Typ stroja M1/M2/M3	Character	Nominal	Kategórie bez prirodzeného poradia.
ID študenta	Numeric alebo Character	Nominal	Číslo je identifikátor, nie meraná veličina.
Výsledok áno/nie	Character alebo Numeric s labels	Nominal	Binárna kategorická odozva.

Najčastejšia chyba: Stĺpec s kódmi 1, 2, 3 JMP automaticky považuje za Continuous. Ak ide o skupiny, zmeňte Modeling Type na Nominal alebo Ordinal.

2.3 Import dát

V JMP: File > Open

JMP dokáže otvoriť súbory JMP, CSV, textové súbory, Excel a ďalšie zdroje. Pri CSV skontrolujte oddeľovač, desatinnú bodku/čiarku, hlavičku a kódovanie. Pri Exceli vyberte správny hárok a skontrolujte, či prvý riadok obsahuje názvy stĺpcov.

1. Po importe prezrite prvých a posledných niekoľko riadkov.
2. Otvorte Column Info pre všetky kľúčové premenné.
3. Spustite Analyze > Distribution na numerické premenné a skontrolujte rozsahy, chýbajúce hodnoty a nepravdepodobné kódy.
4. Uložte čistú pracovnú kópiu ako .jmp.

2.4 Chýbajúce hodnoty a neplatné kódy

Numerická chýbajúca hodnota sa v JMP zobrazuje ako bodka. Prázdny textový údaj je tiež chýbajúci. Kódy ako 999, -1 alebo „N/A“ nie sú automaticky chýbajúce; pred analýzou ich treba správne prekonvertovať. Nikdy nevymieňajte chýbajúce hodnoty za nulu bez vecného dôvodu.

2.5 Vzorce a odvodené stĺpce

V JMP: Cols > New Column > Column Properties > Formula

Formula vytvorí reprodukovateľný odvodený stĺpec. Pri úprave vstupných údajov sa výsledok automaticky prepočíta. Typické príklady: rozdiel po – pred, logaritmus, kategorizácia, centrovanie, štandardizácia a predikčné skóre.

Úloha	Príklad vzorca / princíp
Rozdiel	Skore_po - Skore_pred
Priemer viacerých položiek	Mean(:Polozka1, :Polozka2, :Polozka3)
Logaritmus	Log(:Koncentracia)
Indikátor	If(:Hodnota >= 50, "splnil", "nesplnil")
Centrovanie	Teplota - Col Mean(Teplota)

2.6 Transformácie tabuliek

Cieľ	Cesta	Poznámka
Súhrn podľa skupín	Tables > Summary	Vytvorí novú tabuľku s N, priemerom, SD, min/max a ďalšími štatistikami.
Široký → dlhý formát	Tables > Stack	Viac meracích stĺpcov uloží do stĺpcov Label a Data.
Dlhý → široký formát	Tables > Split	Kategória určí nové stĺpce.
Spojenie tabuliek	Tables > Join	Definujte kľúč a skontrolujte duplicity.
Podmnožina dát	Tables > Subset	Vytvorí novú tabuľku iba z vybraných riadkov/stĺpcov.

Cvičenie 2A – kontrola dát | Dáta: 02_proces_pevnost.csv

Otvorte 02_proces_pevnost.csv. Nastavte Stroj ako Nominal a ostatné analytické premenné ako Continuous. Skontrolujte rozsahy, počet riadkov a chýbajúce hodnoty. Vytvorte nový stĺpec Teplota_c = Teplota – 170 pomocou Formula.

3. Exploračná analýza a vizualizácia

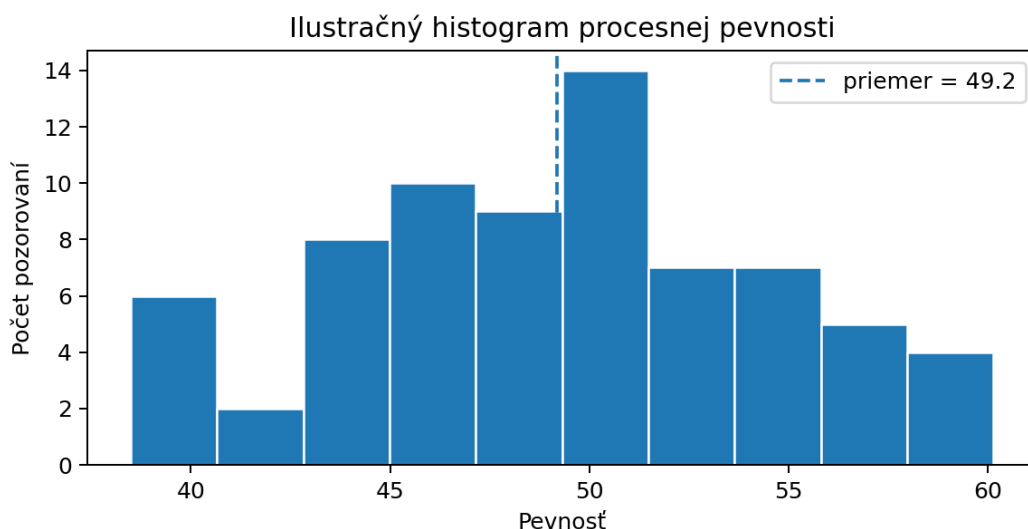
Exploračná analýza (EDA) má predchádzať testu alebo modelu. Cieľom nie je iba „pozrieť sa na dáta“, ale odhaliť rozdelenie, variabilitu, chýbajúce hodnoty, odľahlé body, nelinearity, rozdiely medzi skupinami a možné chyby zberu.

3.1 Jedna premenná: Distribution

V JMP: **Analyze > Distribution**

1. Presuňte numerickú premennú do Y, Columns.
2. Kategorickú premennú môžete pridať do By, čím vytvoríte samostatný výstup pre každú skupinu.
3. Po spustení použite červený trojuholník na zobrazenie normal quantile plot, confidence interval, testov rozdelenia alebo uloženie súhrnov.

Prvok výstupu	Čo sledujeme
Histogram	Tvar, počet módov, šikmosť, medzery a možné chybné hodnoty.
Box plot	Medián, kvartily, rozsah a odľahlé body.
Quantiles	Medián, kvartily a percentily odolné voči extrémom.
Summary Statistics	N, mean, standard deviation, standard error a interval.
Normal Quantile Plot	Približná linearita bodov podporuje normálny model; odchýlky v chvostoch upozorňujú na šikmosť alebo extrémny.



Obrázok 1. Histogram je začiatok analýzy, nie dôkaz normality.

3.2 Graph Builder

V JMP: **Graph > Graph Builder**

Graph Builder je univerzálna platforma na tvorbu grafov pretiahnutím premenných do oblastí X, Y, Overlay, Color, Wrap a Group. Typ grafu zvolíte ikonami nad plochou. Graf zostáva prepojený s tabuľkou: označené body sa zvýraznia vo všetkých otvorených pohľadoch.

Otázka	Odporúčaný graf
Aké je rozdelenie jednej numerickej premennej?	Histogram + box plot.
Líšia sa skupiny?	Box plot alebo violin plot podľa kategórie na X.
Sú dve numerické premenné spojené?	Scatterplot, prípadne smooth alebo fit line.
Ako sa mení hodnota v čase?	Line chart so správne zoradeným časom.
Je prítomná interakcia?	Priemer odozvy podľa faktora X, druhý faktor v Overlay; neparalelné čiary naznačujú interakciu.

3.3 Výber bodov, lasso a row states

Výber v grafe mení stav riadkov. Vybrané riadky môžete označiť farbou, symbolom, skryť alebo vylúčiť. Vylúčenie z analýzy používajte transparentne a iba s vecným odôvodnením; „zlepšenie p-hodnoty“ nie je dôvod na odstránenie bodu.

Upozornenie: Odľahlý bod môže byť chyba, ale aj najdôležitejšie pozorovanie. Najprv overte meranie a protokol, potom urobte citlivostnú analýzu s bodom aj bez neho.

Cvičenie 3A – EDA procesu | *Dáta: 02_proces_pevnost.csv*

Pre 02_proces_pevnost.csv vytvorte: (a) histogram Pevnost, (b) box plot Pevnost podľa Stroj, (c) scatterplot Pevnost vs Teplota s farbou podľa Stroj. Zaznamenajte tri pozorovania, ktoré sú viditeľné v grafoch bez použitia p-hodnoty.

ČASŤ II

Štatistická inferencia

4. Logika testovania a intervalov spoľahlivosti

Štatistická inferencia rozlišuje medzi variabilitou vo výbere a tvrdením o širšej populácii alebo procese. JMP vypočíta testy a intervaly, ale používateľ musí určiť dizajn štúdie, nezávislosť pozorovaní, vhodný model a vecný význam výsledku.

4.1 Od otázky k modelu

Krok	Otázka
1. Cieľ	Opisujeme dáta, porovnávame skupiny, odhadujeme vzťah alebo predikujeme?
2. Odozva Y	Je numerická, binárna, počet, čas do udalosti alebo iný typ?
3. Vysvetľujúce X	Sú kontinuálne, kategorické, fixné, náhodné alebo blokovacie?
4. Dizajn	Sú pozorovania nezávislé, párované, opakované, blokové alebo experimentálne randomizované?
5. Predpoklady	Je model rezíduí a variancie primeraný?
6. Záver	Aký je odhad efektu, interval neistoty, praktický význam a obmedzenie?

4.2 Hypotéza, p-hodnota a alfa

Nulová hypotéza H_0 vyjadruje model bez sledovaného efektu. P-hodnota je pravdepodobnosť výsledku aspoň takého extrémneho, ako pozorovaný, za predpokladu platnosti H_0 a ostatných predpokladov modelu. Nie je to pravdepodobnosť, že H_0 je pravdivá. Hladina α je vopred zvolené pravidlo rozhodovania, nie magická hranica vedeckej pravdy.

Správny zápis: Namiesto „výsledok je významný“ uveďte smer a veľkosť rozdielu, interval spoľahlivosti, p-hodnotu a kontext. Napr.: „Metóda B zvýšila skóre v priemere o 4,7 bodu viac než metóda A; 95 % CI ...; Welchov t-test $p = \dots$ “

4.3 Interval spoľahlivosti a veľkosť efektu

Interval spoľahlivosti vyjadruje rozsah hodnôt kompatibilných s dátami a modelom. Široký interval signalizuje malú presnosť. Veľkosť efektu môže byť rozdiel priemerov v pôvodných jednotkách, štandardizovaný rozdiel, pomer šancí, sklon regresie alebo percentuálna zmena. Pri aplikovanej práci je pôvodná jednotka často najzrozumiteľnejšia.

4.4 Predpoklady: čo skutočne kontrolovať

Predpoklad	Ako ho posúdiť	Poznámka
Nezávislosť	Vyplýva zo spôsobu výberu a randomizácie.	Graf rezíduí ju sám nedokáže potvrdiť.
Správna funkčná forma	Graf Y vs X, residual by predicted, partial plots.	Nelinearitu riešte transformáciou alebo vhodným členom modelu.

Predpoklad	Ako ho posúdiť	Poznámka
Približná normalita rezíduí	Normal quantile plot rezíduí.	Pri veľkých výberoch je dôležitejšia výrazná šikmosť a extrémny než malé odchýlky.
Konštantný rozptyl	Residual by predicted a porovnanie rozptylov skupín.	Pri dvoch skupinách je bezpečnou voľbou Welchov t-test.
Bez dominantných bodov	Leverage, Cook's D, studentized residuals.	Vplyvný bod nie je automaticky chybný.

5. Porovnávanie priemerov

5.1 Jednovýberový t-test

V JMP: **Analyze > Distribution > červený trojuholník pri premennej > Test Mean**

Použite, keď porovnáвате priemer jednej numerickej premennej s referenčnou hodnotou μ_0 . V dialógu zadajte Hypothesized Mean. Reportujte výberový priemer, štandardnú odchýlku, rozdiel od μ_0 , 95 % interval a p-hodnotu.

5.2 Dve nezávislé skupiny

V JMP: **Analyze > Fit Y by X**

1. Numerickú odozvu vložte do Y, Response.
2. Dvojúrovňovú kategorickú premennú vložte do X, Factor.
3. Po spustení: červený trojuholník pri Oneway Analysis > Means/Anova/Pooled t alebo t Test; pri nerovnakých rozptyloch použite Welchovu verziu.
4. Zobrazte box plot, intervaly priemerov a skontrolujte rozdelenia podľa skupín.

Ak je X Nominal a Y Continuous, platforma Fit Y by X vytvorí Oneway Analysis. Pre dve skupiny je test rozdielu priemerov ekvivalentný príslušnému t-testu; pri viacerých skupinách ide o jednofaktorovú ANOVA.

5.3 Párový t-test

V JMP: **Analyze > Matched Pairs**

Použite pri dvoch meraniach na tej istej jednotke alebo pri prirodzene spárovaných jednotkách. Analýza sa týka rozdielov v pároch, nie dvoch nezávislých súborov. Skontrolujte rozdelenie rozdielov a odľahlé páry.

Situácia	Správna metóda
Iní študenti v skupine A a B	Nezávislý t-test / Welch.
Tí istí študenti pred a po kurze	Párový t-test.
Vzorky z dvoch výrobných liniek bez párovania	Nezávislý t-test.
Ľavé a pravé oko tej istej osoby	Párový t-test.

5.4 Neparametrické alternatívy

Pri ordinálnej odozve, výraznej šikmosti, malom výbere alebo odľahlých hodnotách zvážte Wilcoxonov test. V Oneway platforme je dostupný Wilcoxon/Kruskal-Wallis test; pri pároch použite Wilcoxon Signed Rank. Tieto testy nie sú automaticky „testy mediánov“ bez ďalších predpokladov; interpretujú sa ako rozdiel v rozdeleniach alebo typickej polohe.

Cvičenie 5A – metódy výučby | Dáta: 01_studenti_metoda.csv

V 01_studenti_metoda.csv porovnajte premennú Zmena medzi metódami A a B pomocou Fit Y by X. Zobrazte box plot, Welchov t-test a interval rozdielu. Potom vykonajte párový t-test Skore_pred vs Skore_po pre všetkých študentov.

Kontrola výsledku

Priemerná zmena $A \approx 4.35$, $B \approx 9.57$. Welchov test: $t \approx -4.42$, $p \approx 0.0001$. Párový test pre všetkých: $t \approx 10.00$, $p < 0,001$. Záver musí rozlíšiť celkové zlepšenie od rozdielu medzi metódami.

6. ANOVA a viacfaktorové modely

6.1 Jednofaktorová ANOVA

V JMP: Analyze > Fit Y by X

Ak má X viac než dve úrovne a je Nominal/Ordinal, JMP vytvorí Oneway Analysis. ANOVA testuje H_0 : všetky skupinové priemery sú rovnaké. Významný F-test hovorí, že aspoň jeden priemer sa líši; neurčuje, ktoré dvojice sú rozdielne.

1. Zobrazte Means/Anova a súhrn priemerov.
2. Skontrolujte box plot, počet pozorovaní a rozptyly skupín.
3. Použite Means Comparisons, napr. Tukey HSD, ak potrebujete všetky párové porovnania s kontrolou chybovosti.
4. Reportujte F, stupne voľnosti, p, priemery a intervaly; doplňte vecný význam.

Viacnásobné porovnania Nevykonávajte veľký počet obyčajných t-testov bez korekcie. Tukey HSD je vhodný na všetky párové porovnania; Dunnett na porovnanie viacerých skupín s kontrolou.

6.2 Dvojfaktorová ANOVA a interakcia

V JMP: Analyze > Fit Model

1. Odozvu presuňte do Y.
2. Vyberte oba faktory a kliknite Add.
3. Pre interakciu označte oba faktory a kliknite Cross; model obsahuje A, B a $A*B$.
4. Personality: Standard Least Squares; Emphasis: Effect Leverage alebo Minimal Report podľa cieľa.

Interakcia znamená, že efekt jedného faktora závisí od úrovne druhého. Pri významnej alebo vecne veľkej interakcii neinterpretujte hlavné efekty izolovane. Zobrazte interaction plot a porovnávajte jednoduché efekty alebo LSMeans v relevantných rezoch.

6.3 Fixné a náhodné efekty

Fixný faktor reprezentuje konkrétne úrovne záujmu, napr. tri vyučovacie metódy. Náhodný faktor reprezentuje náhodný výber z širšej populácie úrovní, napr. triedy alebo šarže. Zmiešané modely vyžadujú správne označenie Random Effect a často metódu REML. Rozhodnutie je vecné, nie iba softvérové.

Cvičenie 6A – rozdiely medzi strojmi | Dáta: 02_proces_pevnost.csv

V 02_proces_pevnost.csv vykonajte jednofaktorovú ANOVA Pevnost podľa Stroj. Zobrazte Tukey HSD a porovnajte výsledok s box plotom. Potom v Fit Model pridajte Teplota ako kovariátu a posúďte, či rozdiely medzi strojmi pretrvávajú po zohľadnení teploty.

Kontrola výsledku

Jednofaktorová ANOVA: $F \approx 8.78$, $p \approx 0.0004$. Po pridaní teploty sa odhad stroja môže zmeniť, pretože teplota vysvetľuje časť variability. Presné porovnania čítajte z LSMeans/Tukey výstupu.

7. Korelácia a regresia

7.1 Korelácia

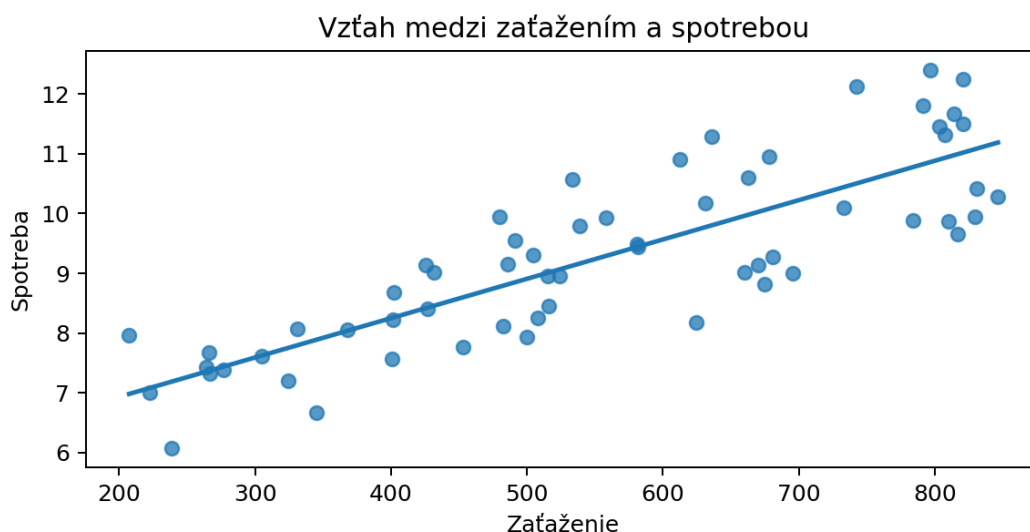
V JMP: Analyze > Multivariate Methods > Multivariate

Pearsonova korelácia r meria lineárnu asociáciu medzi dvoma numerickými premennými. Hodnota blízka ± 1 znamená silný lineárny vzťah, nie nevyhnutne príčinnosť. Pred interpretáciou vždy zobrazte scatterplot; jediný extrémny bod môže koreláciu dramaticky zmeniť.

7.2 Jednoduchá lineárna regresia

V JMP: Analyze > Fit Y by X

Ak sú Y aj X Continuous, platforma vytvorí Bivariate Fit. Z červeného trojuholníka vyberte Fit Line. Regresný model má tvar $Y = \beta_0 + \beta_1 X + \varepsilon$. Sklon β_1 je očakávaná zmena Y pri zvýšení X o jednu jednotku. Intercept β_0 má vecný význam iba vtedy, ak $X = 0$ patrí do relevantného rozsahu.



Obrázok 2. Scatterplot musí predchádzať interpretácii korelácie a regresie.

7.3 Viacnásobná lineárna regresia

V JMP: Analyze > Fit Model

1. Y vložte do role Y.
2. Prediktory vyberte a kliknite Add. Kategorické prediktory musia mať Modeling Type Nominal/Ordinal.
3. Spustite Standard Least Squares.
4. V červenom trojuholníku zapnite Parameter Estimates, Effect Tests, Actual by Predicted, Residual by Predicted, Normal Quantile Plot a Collinearity Diagnostics/VIF.

Výstup	Interpretácia
Summary of Fit	R^2 je podiel variability vysvetlenej modelom; Adjusted R^2 penalizuje zbytočné členy; RMSE je typická chyba predikcie v jednotkách Y.

Výstup	Interpretácia
Analysis of Variance	Celkový F-test porovnáva model s modelom iba s interceptom.
Effect Tests	Test každého efektu pri zohľadnení ostatných členov modelu.
Parameter Estimates	Odhad, štandardná chyba, t-ratio, p a interval pre každý parameter.
VIF	Vysoké hodnoty upozorňujú na multikolinearitu; koeficienty môžu byť nestabilné.
Residual plots	Hľadajte nelinearitu, lievikovitý rozptyl, zhľuky a extrémny.

7.4 Transformácie, polynómy a interakcie

Nelineárny vzťah možno modelovať polynómom, napr. X a X^2 . V Fit Model vyberte premennú a použite Macros > Polynomial to Degree. Pri interakcii $X_1 * X_2$ sa sklon X_1 mení podľa X_2 . Model zachováva hierarchický: ak obsahuje interakciu alebo kvadratický člen, ponechajte príslušné nižšie členy aj pri vyššej p-hodnote.

7.5 Výber modelu

Automatické stepwise metódy môžu byť užitočné pri predikčnom screeningu, ale zvyšujú riziko optimistických p-hodnôt a nestability. Pri inferencii uprednostnite model definovaný otázkou a dizajnom. Pri predikcii použite validačný súbor alebo krížovú validáciu a porovnávajte predikčnú chybu, nie iba R^2 na tréningových dátach.

Cvičenie 7A – model spotreby | Dáta: 03_regresia_spotreba.csv

Otvorte 03_regresia_spotreba.csv. Vytvorte scatterploty Spotreba s každým prediktorom. V Fit Model odhadnite $\text{Spotreba} \sim \text{Rychlost} + \text{Rychlost}^2 + \text{Zataz} + \text{Teplota_okolie}$. Skontrolujte rezíduá, VIF a interpretujte koeficient Zataženie.

Kontrola výsledku

Model má približne $R^2 = 0.834$, adjusted $R^2 = 0.822$, RMSE = 0.63. Odhad pre Zataz ≈ 0.0063 jednotky spotreby na jednu jednotku zataženia pri ostatných premenných konštantných.

8. Kategórické dáta a logistická regresia

8.1 Kontingenčné tabuľky a chí-kvadrát test

V JMP: Analyze > Fit Y by X

Ak sú Y aj X kategórické, JMP vytvorí Contingency Analysis s mozaikovým grafom. Pearsonov chí-kvadrát test posudzuje nezávislosť premenných. Pri malých očakávaných početnostiach použite Fisherov exaktný test alebo zlučujte kategórie iba s vecným odôvodnením.

8.2 Binárna logistická regresia

V JMP: Analyze > Fit Model

Binárna odozva musí byť Nominal s dvoma úrovňami. JMP zvolí Nominal Logistic personality. Modeluje logit pravdepodobnosti udalosti. Exponenciovaný koeficient je odds ratio: násobok zmeny šancí pri zvýšení prediktora o jednu jednotku, pri ostatných premenných konštantných.

Položka	Čo uviesť v reporte
Udalosť	Ktorá úroveň je modelovaná; overte poradie úrovní.
Koeficient / odds ratio	Smer, veľkosť a interval spoľahlivosti.
Celkový model	Likelihood-ratio test a vhodná miera kvality modelu.
Predikcia	ROC/AUC, confusion matrix iba pri zmysluplne zvolenom prahu.
Kalibrácia	Porovnanie predikovaných a pozorovaných pravdepodobností.

Pozor na „pravdepodobnosť“: Odds ratio 2 neznamená dvojnásobnú pravdepodobnosť. Šanca $p/(1-p)$ a pravdepodobnosť p sú odlišné veličiny.

ČASŤ III

Návrh experimentov (DOE)

9. Princípy DOE a pracovný postup

Návrh experimentov je systematický spôsob, ako meniť viac vstupných faktorov naraz a odhadovať ich účinky na jednu alebo viac odoziev. Na rozdiel od prístupu „jeden faktor naraz“ umožňuje DOE odhadovať interakcie a využíva experimentálne zdroje efektívnejšie.

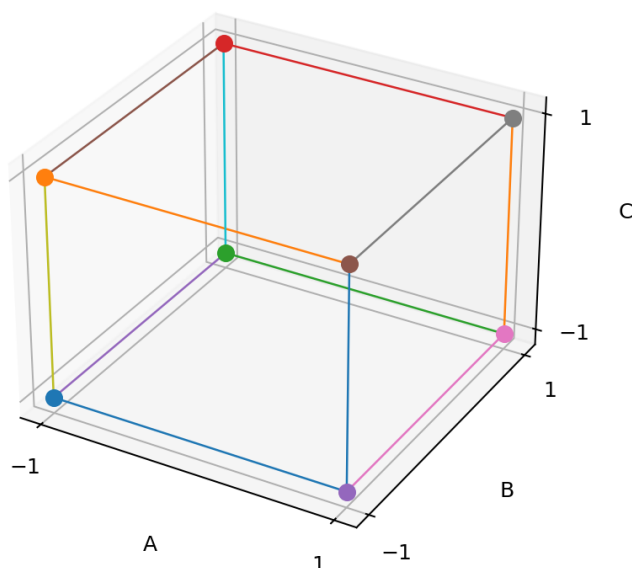
9.1 Základné pojmy

Pojem	Definícia
Odozva (response)	Meraný výsledok experimentu, napr. výťažnosť, pevnosť, čas alebo chybovosť.
Faktor	Riadený vstup alebo podmienka, ktorú experimentátor mení.
Úroveň	Konkrétna hodnota faktora, napr. 60 °C a 80 °C.
Beh (run)	Jedna realizácia kombinácie faktorových úrovní.
Hlavný efekt	Priemerná zmena odozvy medzi úrovňami faktora.
Interakcia	Efekt jedného faktora sa mení podľa úrovne iného faktora.
Replikácia	Nezávislé zopakovanie rovnakého experimentálneho nastavenia.
Stredový bod	Beh na stredných úrovniach kontinuálnych faktorov; pomáha odhaliť zakrivenie.
Blok	Skupina behov vykonaných za podobných rušivých podmienok, napr. deň alebo šarža.
Konfundovanie / aliasing	Dva alebo viac efektov nemožno v danom návrhu oddeliť.

9.2 Tri piliere dobrého experimentu

Princíp	Prečo je dôležitý	Častá chyba
Randomizácia	Chráni odhady pred systematickým časovým alebo priestorovým trendom.	Behy sa vykonávajú v pohodlnom (vzhľadom na realizáciu), ale systematickom poradí.
Replikácia	Poskytuje odhad čistej experimentálnej chyby a zvyšuje presnosť.	Opakované odčítanie tej istej vzorky sa zamieňa za nezávislú replikáciu.
Blokovanie	Oddeľuje známu rušivú variabilitu od efektov faktorov.	Deň alebo šarža sa ignoruje a potom sa zamieňa s faktorom.

Úplný 2^3 faktoriálny plán: 8 vrcholov



Obrázok 3. Pri úplnom experimente 2^3 plán obsahuje všetky kombinácie nízkych (-1) a vysokých ($+1$) úrovní.

9.3 Odporúčaná postupnosť krokov DOE v JMP

1. Define: sformulujte cieľ, odozvy, faktory, rozsahy, obmedzenia a minimálny prakticky dôležitý efekt.
2. Design: vyberte typ návrhu, model, počet behov, randomizáciu, bloky, replikácie a stredové body.
3. Evaluate: skontrolujte aliasing, korelácie odhadov, power, prediction variance a uskutočniteľnosť behov.
4. Run: vykonajte experiment podľa Run Order a zapisujte odchýlky od protokolu.
5. Analyze: použite uložený Model script, posúďte efekty, interakcie, rezíduá a lack of fit.
6. Optimize and confirm: použite profiler, zvolte nastavenie a vykonajte nezávislé potvrdzujúce behy.

Zásadné pravidlo: Dizajn sa posudzuje podľa modelu, ktorý má podporovať. Počet behov sám o sebe nehovorí, či je návrh vhodný.

9.4 Kódované a prirodzené jednotky

Dvojúrovňové kontinuálne faktory sa často kódujú ako -1 a $+1$. Kódovanie uľahčuje porovnávanie efektov rôznych faktorov. Pri reportovaní však uveďte aj prirodzené jednotky. Pre stred m a polovičný rozsah intervalu faktora d platí $x_{\text{kód}} = (x - m)/d$.

10. Úplný faktoriálny plán

10.1 Kedy ho použiť

Úplný faktorový/faktoriálny plán obsahuje všetky kombinácie úrovní faktorov. Je vhodný pri malom počte faktorov, keď potrebujete odhadnúť hlavné efekty aj interakcie bez aliasingu. Pre k dvojúrovňových faktorov má 2^k základných behov; ich počet rýchlo rastie.

Počet faktorov k	Základné behy 2^k
2	4
3	8

Počet faktorov k	Základné behy 2^k
4	16
5	32
6	64

10.2 Vytvorenie návrhu v JMP

V JMP: DOE > Classical > Full Factorial Design

1. V časti Responses pomenujte odozvu, nastavte Goal a prípadne Lower/Upper Limit.
2. Pridajte Continuous alebo Categorical factors a zadajte nízke/vysoké úrovne.
3. Kliknite Continue a nastavte randomizáciu, počet replikácií a stredových bodov.
4. Kliknite Make Table. JMP vytvorí tabuľku a uloží skripty Design a Model.
5. Po vykonaní behov zapisujte výsledky výhradne do príslušných riadkov podľa Run Order.

10.3 Analýza úplného faktoriálneho plánu

V JMP: V tabuľke návrhu spustíte skript Model; alternatívne Analyze > Fit Model

Začnite modelom, pre ktorý bol návrh vytvorený. Pri 2^3 pláne môže úplný model obsahovať A, B, C, AB, AC, BC a ABC. Pri malom nereplikovanom pláne však úplný model nemá reziduálne stupne voľnosti. Vtedy treba mať vopred premyslenú stratégiu: replikáciu, stredové body, predpoklad zanedbateľnosti vyšších interakcií alebo vhodný screeningový prístup.

10.4 Hlavný efekt a interakcia

V dvojúrovňovom pláne je efekt A rozdiel priemernej odozvy pri $A = +1$ a $A = -1$. Interakcia AB porovnáva, či sa efekt A mení medzi úrovňami B. V grafe interakcie sú paralelné čiary znakom malej interakcie; pretínajúce alebo výrazne neparalelné čiary signalizujú interakciu.

Cvičenie 10A – filtrácia | *Dáta: 04_doe_filtracia.csv*

Otvorte 04_doe_filtracia.csv. V Fit Model použite iba faktoriálne body A_kod, B_kod, C_kod (stredové body dočasne vylúčte) a model A|B|C do stupňa 2, teda hlavné efekty a dvojfaktorové interakcie. Zobrazte Effect Summary, Parameter Estimates, Actual by Predicted a Profiler.

Kontrola výsledku

V generovaných dátach sú najsilnejšie hlavné efekty A a B a interakcia A*B. Odhad koeficientov modelu s úplnou 3-faktorovou interakciou: $A \approx 6.29$, $B \approx 3.71$, $C \approx -3.21$, $A*B \approx 3.26$. Koeficient v kódovanom modeli je polovica príslušného efektu.

11. Frakčný faktoriálny a screeningový plán

11.1 Prečo frakcia

Pri väčšom počte faktorov je úplný plán príliš rozsiahly. Frakčný faktoriálny plán používa iba časť kombinácií a zámerne konfunduje niektoré efekty. Je určený najmä na screening: identifikáciu malého počtu aktívnych faktorov medzi mnohými kandidátmi.

11.2 Rozlíšenie (resolution)

Rozlíšenie	Typická interpretácia
III	Hlavné efekty môžu byť aliasované s dvojfaktorovými interakciami. Vhodné len pri silnom predpoklade malých interakcií.
IV	Hlavné efekty nie sú aliasované s dvojfaktorovými interakciami, ale dvojfaktorové interakcie môžu byť aliasované navzájom.
V	Hlavné efekty a dvojfaktorové interakcie sú navzájom oddeliteľné; dvojfaktorové interakcie môžu byť aliasované s trojfaktorovými.

11.3 Vytvorenie klasického screeningového návrhu

V JMP: DOE > Classical > Two Level Screening > Screening Design

1. Definujte odozvy a faktory s nízkymi/vysokými úrovňami.
2. Vyberte návrh zo zoznamu podľa počtu behov, rozlíšenia a aliasovej štruktúry.
3. Pridajte stredové body pre kontinuálne faktory, ak chcete testovať zakrivenie.
4. Randomizujte run order alebo definujte bloky, ak randomizácia nie je úplne možná.
5. Pred Make Table prezrite aliasing a podporovaný model.

11.4 Definitive Screening Design (DSD)

V JMP: DOE > Definitive Screening > Definitive Screening Design

DSD je moderný screeningový návrh, ktorý pri kontinuálnych faktoroch používa tri úrovne a umožňuje odhadovať hlavné efekty bez konfundovania s dvojfaktorovými interakciami; poskytuje aj informáciu o zakrivení. Je výhodný pri mnohých faktoroch a očakávaní, že aktívnych bude iba niekoľko. Analýza musí rešpektovať princíp sparsity a hierarchiu modelu.

Sekvenčné experimentovanie: Screening nie je posledný krok. Po výbere faktorov nasleduje zúženie rozsahov, doplnenie návrhu alebo samostatný RSM experiment a nakoniec potvrdzujúce behy.

11.5 Nereplikované screeningové plány

Ak chýba čistý odhad chyby, JMP môže ponúknuť metódy založené na sparsity efektov, napr. efektové grafy alebo špecializované screeningové analýzy. Výsledky vnímajte ako výber kandidátov, nie ako definitívny dôkaz. Silné interakcie, časový trend alebo zmena variability môžu viesť k nesprávnemu výberu.

12. Custom Design

12.1 Kedy použiť Custom Design

Custom Design vytvára optimálne návrhy pre konkrétny model a obmedzenia. Je vhodný pri kombinácii kontinuálnych a kategorických faktorov, zakázaných kombináciách, nerovnakých nákladoch, blokoch, split-plot štruktúre alebo presne danom počte behov.

V JMP: DOE > Custom Design

12.2 Základný postup

1. Responses: pomenujte každú odozvu a nastavte cieľ (Maximize, Minimize, Match Target, None).
2. Factors: pridajte Continuous, Categorical, Discrete Numeric, Mixture alebo ďalší vhodný typ; zadajte rozsahy.

3. Model: pridajte hlavné efekty, interakcie a kvadratické členy, ktoré chcete odhadovať. Použite RSM macro, ak navrhujete kvadratický model.
4. Number of Runs: nechajte odporúčaný počet alebo nastavte rozpočet; počet musí poskytovať dostatok stupňov voľnosti a presnosti.
5. Design Generation: vygenerujte návrh, prípadne niekoľkokrát s iným random seed a porovnajte kvalitu.
6. Design Evaluation: skontrolujte power, prediction variance, correlations of estimates a design diagnostics.
7. Make Table a uložte všetky skripty.

12.3 D-optimalita a I-optimalita

Kritérium	Zameranie	Typické použitie
D-optimal	Presnosť odhadov parametrov modelu; maximalizuje informačný determinant.	Screening a inferencia o efektoch.
I-optimal	Priemerná predikčná variancia v experimentálnom priestore.	Modelovanie odozvovej plochy a optimalizácia.

12.4 Obmedzenia a uskutočniteľnosť

Zakázané kombinácie definujte pred generovaním návrhu. Po vytvorení návrhu svojvoľne nemažte „nepohodlné“ behy; zničili by ste optimalitu a možno aj odhadnuteľnosť modelu. Ak sa experiment zmení, návrh augmentujte alebo prepočítajte.

12.5 Power a minimálne dôležitý efekt

Power analýza DOE vyžaduje odhad variability a veľkosti efektu, ktorý má experiment spoľahlivo zachytiť. Nežadavajte nerealisticky malú chybu iba preto, aby postačoval menší návrh. Pri neistote urobte pilotné replikácie alebo konzervatívny scenár.

13. Plocha odozvy a optimalizácia

13.1 Kedy prejsť na RSM

Response Surface Methodology (RSM) sa používa po identifikácii dôležitých kvantitatívnych faktorov, keď je cieľom nájsť optimum, cieľovú hodnotu alebo robustnú oblasť. Kvadratický model obsahuje lineárne členy, interakcie a druhé mocniny. Dvojúrovňový faktoriálny plán sám osebe nevie oddeliť všeobecné zakrivenie od konkrétnych kvadratických efektov.

13.2 Vytvorenie klasického RSM návrhu

V JMP: DOE > Classical > Response Surface Design

Typické návrhy sú central composite design (CCD) a Box–Behnken design (BBD). Zvoľte faktory, rozsahy, počet stredových bodov a prípadné bloky. Pri fyzikálnych obmedzeniach a zmiešaných faktoroch je často vhodnejší DOE > Custom Design s RSM modelom.

13.3 Kvadratický model

Pre dva faktory má model tvar $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{12}X_1X_2 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \varepsilon$. Lineárne členy opisujú sklon, interakcia natočenie plochy a kvadratické členy zakrivenie. Model interpretujte predovšetkým graficky pomocou profilov a contour plotov.



Obrázok 4. Kontúry ukazujú kombinácie faktorov s podobnou predikovanou odozvou; body označujú experimentálne behy.

13.4 Prediction Profiler

V JMP: Fit Model report > červený trojuholník > Factor Profiling > Profiler

Profiler zobrazuje predikovanú odozvu pri zvolených nastaveniach faktorov a lokálny sklon. Pri viacerých odozvách nastavte desirability functions a použite Maximize Desirability. Výsledok je modelové odporúčanie; musí byť overené potvrdzujúcimi behmi.

13.5 Desirability

Cieľ	Nastavenie
Maximalizovať	Definujte nízku neprijateľnú a vysokú žiaducu hodnotu.
Minimalizovať	Definujte vysokú neprijateľnú a nízku žiaducu hodnotu.
Trafiť cieľ	Zadajte target a tolerované dolné/horné hranice.
Iba obmedzenie	Odozva môže mať nízku dôležitosť, ale musí zostať v povolenom pásme.

Cvičenie 13A – plocha odozvy | Dáta: 05_doe_rsm.csv

V 05_doe_rsm.csv odhadnite kvadratický model pre Čistota s X1_kod, X2_kod, ich interakciou a druhými mocninami. Zobrazte Profiler a Contour Profiler. Nájdite nastavenie s vysokou predikovanou čistotou a navrhните dva potvrdzujúce behy.

Kontrola výsledku

Kvadratický model má približne $R^2 = 0.990$. Očakávajte kladné lineárne členy, záporné kvadratické členy a mierne kladnú interakciu. Optimum leží vo vnútri experimentálneho priestoru, nie na extrémnom rohu; presnú polohu určte profilerom.

14. Analýza, diagnostika a reporting DOE

14.1 Začnite uloženým skriptom Model

Tabuľka vytvorená DOE platformou obsahuje skript Model, ktorý prenesie správne modelové členy a role. Je bezpečnejší než ručné zostavenie modelu, najmä pri blokoch, aliasoch, zmesiach alebo split-plot štruktúre.

14.2 Effect Summary a hierarchia

Effect Summary zoradzuje efekty podľa $\text{LogWorth} = -\log_{10}(p)$. P-hodnota je iba jeden signál. Pri redukcii modelu zachovajte hierarchiu: významná interakcia AB vyžaduje A aj B; kvadratický člen A^2 vyžaduje A. V DOE navyše rešpektujte aliasing a pôvodný cieľ návrhu.

14.3 Lack of fit a čistá chyba

Ak máte replikované nastavenia, reziduálna chyba sa môže rozdeliť na pure error a lack of fit. Významný lack of fit znamená, že model nedostatočne opisuje systematickú štruktúru v experimentálnom priestore. Riešením môže byť transformácia, kvadratické členy, interakcie, užší rozsah alebo nový návrh; nie slepé odstránenie bodov.

14.4 Diagnostika rezíduí

Graf / ukazovateľ	Čo hľadať	Možný krok
Residual by Predicted	Lievik, oblúk, skupiny.	Transformácia odozvy, ďalší člen modelu, model rozptylu.
Residual by Row	Časový trend alebo skok.	Blok, trendový faktor, kontrola procesu a poradia behov.
Normal Quantile Plot	Extrémne chvosty alebo viac populácií.	Overiť meranie; robustná/citlivostná analýza.
Studentized residual	Veľké absolútne hodnoty.	Overiť protokol a vplyv na záver.
Cook's D / leverage	Dominantné experimentálne body.	Pri optimálnych DOE môže byť leverage zámerne nerovnomerný; neodstraňovať automaticky.

14.5 Potvrdzujúce behy

Po optimalizácii vykonajte nové behy pri odporúčanom nastavení, ideálne v inom čase alebo šarži. Porovnajte pozorovaný výsledok s predikčným intervalom, nie iba s bodovou predikciou. Ak optimum leží na hranici priestoru, zväžte bezpečné rozšírenie rozsahu a nový sekvenčný experiment.

14.6 Šablóna reportu DOE

Časť	Minimálny obsah
Cieľ	Odozvy, faktory, rozsahy, prakticky dôležitý efekt.
Návrh	Typ, počet behov, randomizácia, replikácie, stredové body, bloky a obmedzenia.
Realizácia	Odchýlky od run order, zlyhané behy, merací systém.
Model	Modelové členy, kódovanie, metóda odhadu a redukcie.
Výsledky	Odhady efektov a intervaly, interakcie, grafy a diagnostika.

Časť	Minimálny obsah
Optimalizácia	Nastavenia, desirability, predikcia a interval.
Potvrdenie	Nezávislé potvrdzujúce behy a rozdiel oproti predikcii.
Obmedzenia	Rozsah platnosti, neistota, nepreskúmané faktory.

ČASŤ IV

Cvičenia, riešenia a rýchle pomôcky

15. Súhrnné cvičenia s riešeniami

15.1 Mini-projekt zo štatistiky

Zadanie | *Dáta: 02_proces_pevnost.csv*

Použite 02_proces_pevnost.csv. Formulujte výskumnú otázku o vplyve stroja, teploty a vlhkosti na pevnosť. Urobte EDA, odhadnite vhodný regresný/ANOVA model, skontrolujte interakciu Stroj*Teplota, diagnostikujte model a napíšte záver do 150 slov. Povinne uveďte jednu vecnú neistotu a jeden návrh na zlepšenie zberu dát.

Odporúčaná osnova riešenia

1. Nastaviť Stroj ako Nominal. 2. Grafy Pevnosť podľa stroja a proti kontinuálnym prediktorom. 3. Fit Model s hlavnými efektmi a vopred odôvodnenou interakciou. 4. Effect Tests, Parameter Estimates/LSMeans, residual plots a vplyvné body. 5. Záver formulovať v jednotkách pevnosti a zdôrazniť, že pozorovacie dáta samy osebe nedokazujú kauzalitu.

15.2 Mini-projekt DOE

Zadanie | *Dáta: 04_doe_filtracia.csv*

Použite 04_doe_filtracia.csv. Zrekonštruujte cieľ experimentu, identifikujte faktory a úrovne, analyzujte hlavné efekty a interakcie, posúďte stredové body ako signál zakrivenia a pomocou profileru navrhnete nastavenie na maximalizáciu výťažnosti. Uveďte plán troch potvrdzujúcich behov.

Odporúčaná osnova riešenia

1. Rozlíšiť 16 faktoriálnych bodov a 3 stredové body. 2. Odhadnúť hierarchický model s hlavnými efektmi a relevantnými dvojfaktorovými interakciami. 3. Silný pozitívny efekt teploty, pozitívny efekt času, negatívny efekt pH a pozitívna interakcia teplota*čas. 4. Overiť rezíduá a rozdiel stredových bodov oproti lineárnemu modelu. 5. Potvrdenie pri odporúčanom bode a aspoň pri jednom blízkom robustnom bode.

15.3 Kontrolné otázky

1. Aký je rozdiel medzi Data Type a Modeling Type?
2. Prečo významný ANOVA F-test nestačí na tvrdenie, ktoré konkrétne skupiny sa líšia?
3. Kedy je Welchov t-test vhodnejší než pooled t-test?
4. Čo znamená významná interakcia A*B?
5. Prečo korelácia sama osebe nepreukazuje príčinnosť?
6. Aký je rozdiel medzi replikáciou a opakovaným meraním tej istej vzorky?
7. Čo získame pridaním stredových bodov do dvojúrovňového plánu?
8. Čo znamená aliasing vo frakčnom faktoriálnom návrhu?
9. Prečo musí byť model pri interakciách hierarchický?
10. Prečo je po optimalizácii potrebný potvrdzujúci experiment?

16. Najčastejšie chyby

Chyba	Dôsledok	Náprava
ID alebo kód skupiny je Continuous	JMP odhaduje nezmyselný lineárny sklon.	Nastaviť Nominal/Ordinal.
Test bez grafu	Prehliadnutá nelinearita, extrém alebo zmes populácií.	EDA pred testom.
Párové dáta analyzované ako nezávislé	Strata informácie alebo nesprávna chyba.	Matched Pairs / model opakovaných meraní.
Mnoho t-testov bez korekcie	Nafúknutá pravdepodobnosť falošných nálezov.	Tukey, Dunnett alebo plánované kontrasty.
Interpretácia $p > 0,05$ ako „žiadny efekt“	Zámena nedostatku dôkazu za dôkaz nulového efektu.	Pozrieť odhad a interval, posúdiť power.
Odstránenie bodu len preto, že je odľahlý	Skreslený výsledok a selektívna analýza.	Overiť meranie, zdokumentovať a urobiť citlivostnú analýzu.
DOE bez randomizácie	Efekt faktora sa môže zamieňať s časom.	Randomizovať alebo blokovať.
Plán „jeden faktor naraz“	Interakcie sa nedajú zistiť, experiment je neefektívny.	Použiť faktoriálny DOE.
Vymazanie behov z optimálneho návrhu	Strata odhadnuteľnosti alebo presnosti.	Augmentovať/prepočítať návrh.
Optimalizácia bez potvrdenia	Modelové optimum môže byť optimistické alebo lokálne.	Nové potvrdzujúce behy.

17. Rýchla mapa menu

Úloha	Cesta v JMP
Jedna numerická premenná	Analyze > Distribution
Dve premenné / t-test / ANOVA / jednoduchá regresia / kontingenčná tabuľka	Analyze > Fit Y by X
Viacnásobná regresia, viacfaktorová ANOVA, logistická regresia	Analyze > Fit Model
Párové merania	Analyze > Matched Pairs
Korelačná matica	Analyze > Multivariate Methods > Multivariate
Interaktívne grafy	Graph > Graph Builder
Súhrnná tabuľka	Tables > Summary
Široké dáta na dlhé	Tables > Stack
Dlhé dáta na široké	Tables > Split

Úloha	Cesta v JMP
Spojenie tabuliek	Tables > Join
Úplný faktoriálny DOE	DOE > Classical > Full Factorial Design
Frakčný dvojúrovňový screening	DOE > Classical > Two Level Screening > Screening Design
Definitive Screening Design	DOE > Definitive Screening > Definitive Screening Design
Optimálny návrh	DOE > Custom Design
Klasický plán plochy odozvy	DOE > Classical > Response Surface Design
Porovnanie návrhov	DOE > Design Diagnostics > Compare Designs

17.1 Šablóna analýzy

Veta	Obsah
1	Cieľ a dizajn: čo sa zisťovalo a ako boli dáta získané.
2	Dáta: odozva, prediktory, veľkosť výberu, chýbajúce hodnoty.
3	Metóda: použitý model/test a jeho hlavné predpoklady.
4	Výsledok: odhad efektu v jednotkách, 95 % interval, p-hodnota podľa potreby.
5	Diagnostika: či sa objavili závažné problémy.
6	Vecný záver a obmedzenie: čo možno a nemožno tvrdiť.

18. Slovník pojmov a zdroje

Pojem	Význam
Adjusted R ²	R ² upravené o počet parametrov; vhodnejšie na porovnanie vnorených modelov rôznej zložitosti.
Aliasing	Neoddeliteľnosť efektov v neúplnom návrhu.
ANOVA	Analýza rozptylu; porovnáva variabilitu vysvetlenú modelom s reziduálnou variabilitou.
Confidence interval	Interval hodnôt kompatibilných s dátami a modelom pri danej úrovni spoľahlivosti.
Desirability	Funkcia 0–1 vyjadrujúca žiaducosť odozvy; pri viacerých odozvách sa kombinuje.
Effect	Zmena odozvy spojená so zmenou faktora.
Fixed effect	Efekt konkrétnych úrovní, ktoré sú predmetom záujmu.
Interaction	Situácia, keď efekt jedného faktora závisí od iného faktora.

Pojem	Význam
Lack of fit	Časť reziduálnej variability pripísateľná nedostatočnej forme modelu nad rámec čistej chyby.
Least squares means / LSMeans	Modelom upravené priemery, najmä pri nevyvážených dátach alebo kovariátach.
Nominal / Ordinal / Continuous	Modelovacie typy bez poradia / s poradím / numerická spojitá škála.
Profiler	Interaktívny graf predikovanej odozvy podľa nastavení faktorov.
Pure error	Variabilita medzi nezávislými replikáciami pri rovnakom nastavení.
Random effect	Efekt úrovni považovaných za náhodný výber zo širšej populácie.
Residual	Pozorovaná hodnota mínus hodnota predikovaná modelom.
RMSE	Odhad typickej veľkosti reziduálnej chyby v jednotkách odozvy.
VIF	Ukazovateľ kolinearít prediktora s ostatnými prediktormi.

18.1 Oficiálne zdroje JMP

- [JMP Student Edition – informácie a prístup](#)
- [Academic Licensing for Students](#)
- [JMP Documentation](#)
- [JMP 19.1 Online Help – Introduction to DOE](#)
- [Design of Experiments Guide for JMP 19](#)
- [Fitting Linear Models](#)
- [JMP Learning Library](#)
- [Statistics Knowledge Portal – Design of Experiments](#)
- [DOE Full Factorial Design – step-by-step guide](#)
- [DOE Custom Designs – step-by-step guide](#)

18.2 Odporúčaná literatúra

- Montgomery, D. C. Design and Analysis of Experiments. Wiley.
- Box, G. E. P., Hunter, J. S., Hunter, W. G. Statistics for Experimenters. Wiley.
- Kutner, M. H. et al. Applied Linear Statistical Models. McGraw-Hill/Irwin.
- Wasserstein, R. L., Schirm, A. L., Lazar, N. A. Moving to a World Beyond “ $p < 0.05$ ”. The American Statistician.

Záverečný checklist študenta: Je modelovací typ správny? Vidím dáta v grafe? Zodpovedá metóda dizajnu? Uviedol som odhad a interval? Skontroloval som diagnostiku? Je záver v jednotkách problému? Pri DOE som dodržal randomizáciu a navrhol potvrdenie?

Príloha A – inštalácia a aktivácia Student Edition

Aktuálny prístup k Student Edition sa získava cez stránku JMP Student Edition. Používateľ sa prihlási alebo vytvorí účet MyJMP, overí prísľušnosť k akademickej inštitúcii, stiahne verziu pre Windows alebo macOS, nainštaluje ju a pri prvom spustení aktivuje prihlásením. Overenie spravidla vyžaduje akademickú e-mailovú adresu alebo iný doklad oprávnenosti. Licenciu je potrebné po uplynutí obdobia bezplatne obnoviť, pokiaľ používateľ zostáva oprávnený.

Ochrana účtu Prihlasovacie údaje do univerzitného systému zadávajte iba na autentifikačnej stránke vašej inštitúcie alebo v oficiálnom toku MyJMP. JMP podľa aktuálneho inšalačného návodu nedostáva univerzitné prihlasovacie údaje.

Príloha B – obsah dátového balíka

Súbor	Obsah	Kapitoly
00_popis_suborov.csv	Dátový slovník balíka	všetky
01_studenti_metoda.csv	Skóre pred/po a dve vyučovacie metódy	5, 15
02_proces_pevnost.csv	Pevnosť podľa stroja, teploty a vlhkosti	2, 3, 6, 15
03_regresia_spotreba.csv	Spotreba podľa rýchlosti, zaťaženia a teploty	7
04_doe_filtracia.csv	2 ³ DOE s replikáciami a stredovými bodmi	10, 14, 15
05_doe_rsm.csv	CCD-like plán pre dve kontinuálne premenné	13

Poznámka k dátam

Cvičné dáta boli generované výlučne na pedagogické účely. Pri importovaní CSV v prostredí s desatinnou čiarkou skontrolujte, že JMP správne rozpoznal desatinnú bodku a oddeľovač čiarkou.