

LICENČNÍ STUDIUM

GALILEO

STATISTICKÉ ZPRACOVÁNÍ DAT

SEMESTRÁLNÍ PRÁCE

**Statistická analýza jednorozměrných dat
a
ANOVA**

Vladimír Pata

Univerzita Tomáše Bati ve Zlíně

Fakulta technologická

Ústav výrobních technologií

Vavrečkova 275

760 01 Zlín

Česká republika

Obsah

Statistická analýza jednorozměrných dat

Úloha 1. Statistická analýza velkých výběrů	3
Úloha 2. Statistická analýza malých výběrů dle Horna	12
Úloha 3. Statistické testování Test správnosti	15
Úloha 4. Test shodnosti	24
Úloha 5. Párový test	36

ANOVA

Úloha 1. Jednofaktorová ANOVA bez transformace	39
Úloha 2. Jednofaktorová ANOVA s transformací	47
Úloha 3. Dvoufaktorová ANOVA bez opakování	53
Úloha 4. Dvoufaktorová ANOVA s opakováním vyvážená	56

Statistická analýza jednorozměrných dat

Zadání:

Při hodnocení jakosti povrchu reflektoru vyrobeného z polykarbonátu, vyrobeného technikou vstřikování do formy, byl sledován amplitudový parametr drsnosti Rz dle ISO 4287. Vyčíslete charakteristiky šikmosti, špičatosti, nalezněte tvar rozdělení parametru Rz, posuďte zda jsou ve výběru odlehlé hodnoty a pokud ano, využijte transformací pro nalezení nejlepšího odhadu střední hodnoty.

Naměřená data:

Tab. 1.1 Naměřené hodnoty amplitudového parametru drsnosti povrchu Rz (um)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2.86	2.85	2.91	4.23	1.97	1.94	2.34	2.87	2.12	2
(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(18)	(20)
2.14	1.78	2.41	1.86	1.89	2.09	2.26	1.52	2.45	2.79
(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
2.16	1.45	2.24	2.98	2.93	2.92	5.24	4.15	2.91	3.65

(1) Pořadové číslo měření 2.86 Hodnota amplitudového parametru drsnosti povrchu Rz (um)

Užitý program: Adstat125-Student

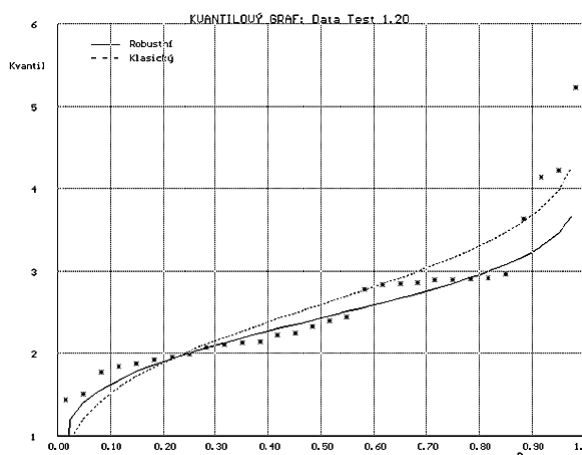
Řešení:

I Průzkumová (explorativní) analýza dat

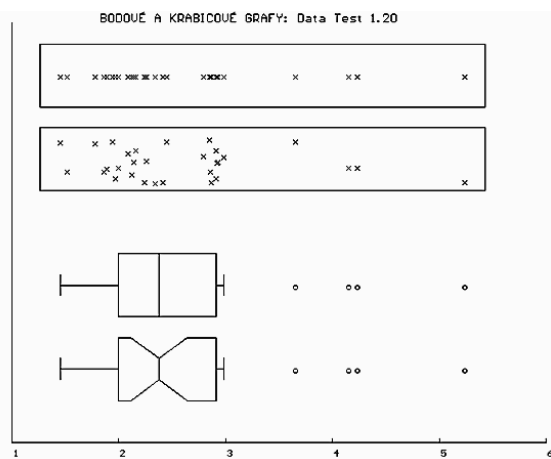
Odhalení stupně symetrie a špičatosti rozdělení

Mezi základní statistické zvláštnosti patří stupeň symetrie a špičatost rozdělení výběru, lokální koncentrace dat a přítomnost vybočujících hodnot (1).

Obr.1.1 Kvantilový graf



Obr 1.2 Bodový a krabicový graf



Kvantilový graf (viz Obr.1.1) ukazuje na nesympetrii naměřených dat a dále detekuje 3 až 4 odlehlé body v pravé části grafu.

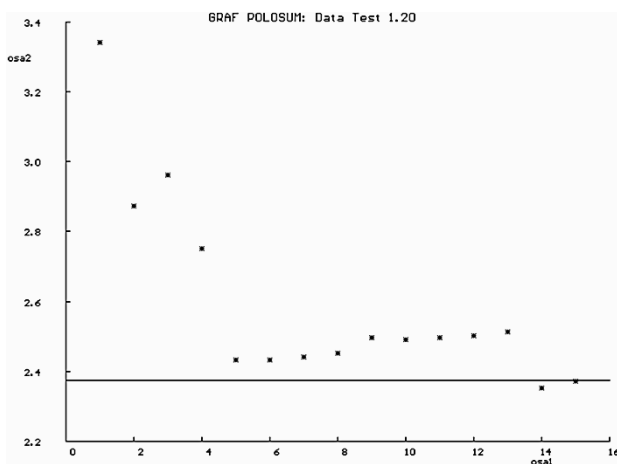
Diagram rozptýlení a Rozmítnutý diagram rozptýlení (viz Obr.1.2 horní část obrázku) detekuje výraznou asymetrii a přítomnost několika, konkrétně čtyřech vybočujících hodnot

Krabicový graf (viz Obr.1.2 dolní část obrázku) ukazuje na nesymetrii naměřených dat, protože polohy $B_{PD}-F_D$ a F_D-M jsou asymetrické oproti polohám $M-F_H$ a $F_H -B_{PH}$.

Na vrubovém diagramu můžeme též specifikovat poměrně široký interval spolehlivosti pro odhad mediánu M , určený I_D a I_H , který je určen s $1-\alpha=0.95$.

Z Kvantilového grafu, a dále grafů rozptýlení a krabicového grafu se shodně jeví, že v souboru měření hodnot R_z se vyskytují vybočující hodnoty (tzv. odlehlá data). Ty však nemohou být vyřazeny ze souboru, protože specifikují „nejhrubější“ část sledovaného povrchu reflektoru, který je rozhodující pro kvalitu povrchu. Zákazníka vždy zajímá nikoliv nejjemnější, ale nejhrubější část povrchu světlometu. Tam totiž dochází k největšímu rozptylu světelných paprsků.

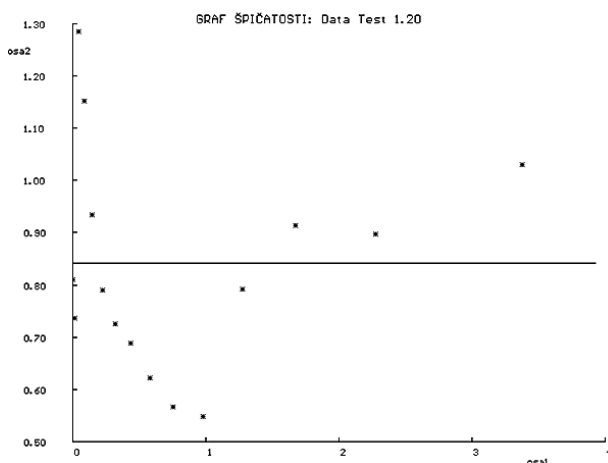
Obr 1.3 Graf polosum



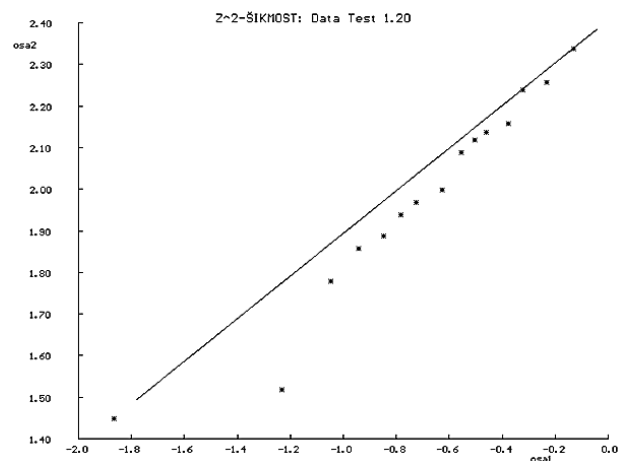
V Graf polosum (viz.Obr.1.3) je patrná systematická odchylka od symetrického rozdělení.

Dále je možné specifikovat, že zoom (rozptyl) bodů na grafu v ose y je relativně velký (3.4 um- 2.2um), což opět naznačuje asymetrii naměřených dat.

Obr 1.4 Graf špičatosti

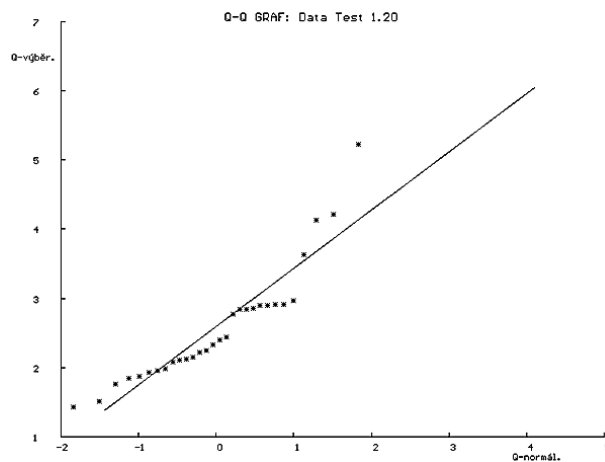


Obr 1.5

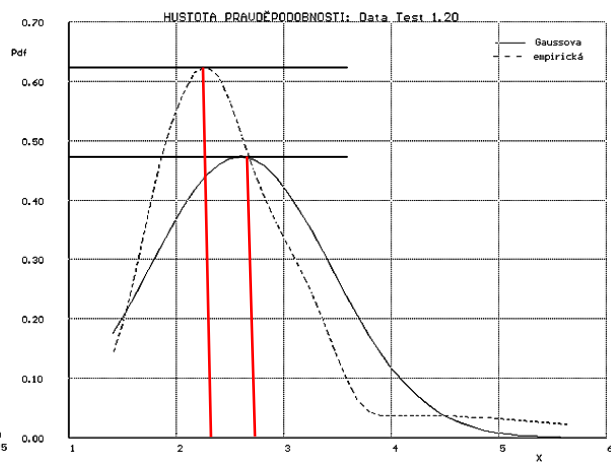


Z Grafu špičatosti (viz.Obr.1.4) je možno určit, že většina bodů neleží rovnoběžně s osou x. Tedy se jedná o asymetricky rozložená data.

Obr 1.6 Q-Q graf



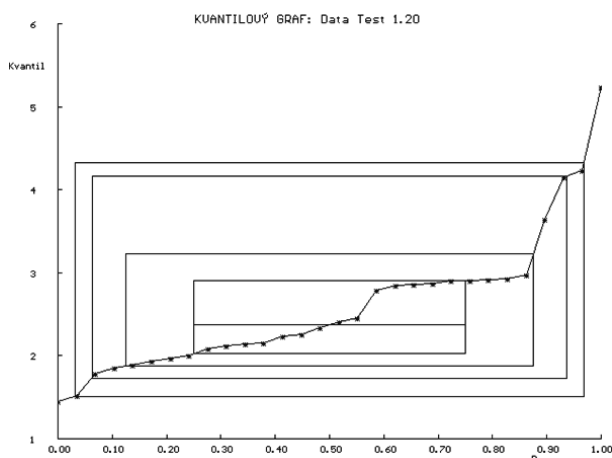
Obr 1.7 Odhad hustoty



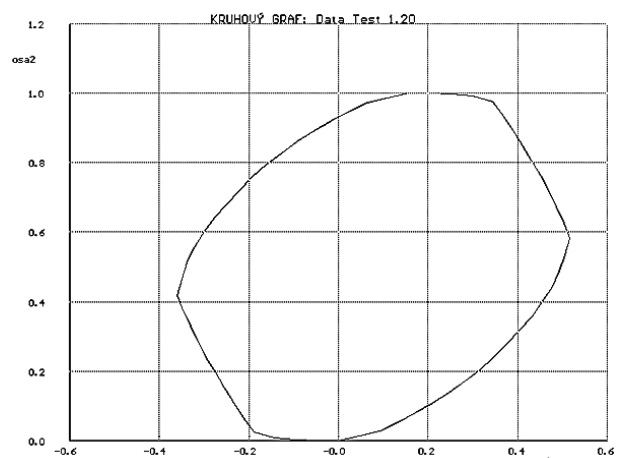
Při posuzování dat pomocí Q-Q grafu je možno konstatovat (viz. Obr.1.6), že většina naměřených dat neleží na přímce, která indikuje teoretické normální rozdělení, a proto předpokládáme asymetrické rozdělení naměřených dat. V tomto případě se zdá, že rozdělení dat je odlišné od rozdělení Normálního a navíc jsou v datech vybočující hodnoty.

Graf jádrového odhadu hustoty pravděpodobnosti (nazývaný též jádrový odhad frekvenční funkce (viz. Obr.1.7) detekuje, že tvarově rozdělení je odlišné od rozdělení Normálního. Průsečík kolmice spuštěné z vrcholové tečny Gaussovy křivky určuje na ose x hodnotu odhadu průměru, průsečík kolmice spuštěné z vrcholové tečny z vrcholu empirické křivky určuje na ose x hodnotu odhadu mediánu. V Obr.1.7 je zcela zřetelně vidět, že odhad průměru je posunut (v důsledku vybočujících hodnot) vůdči odhadu mediánu. Můžeme tedy uvažovat o asymetrii dat, které jsou odlišné od rozdělení Normálního.

Obr 1.8 Graf rozptýlení s kvantily



Obr 1.9 Kruhový graf



U grafu Rozptýlení s kvantily (viz Obr.1.8) nemá kvantilová funkce sigmoidální tvar a je konvexně rostoucí. Tedy vzdálenosti dolních hran kvartilového obdélníku F, Oktilového obdélníku E a

Sedecilového obdélníku D je výrazně kratší, než vzdálenosti mezi jejich horními hranami. Jedná se tedy o zešikmení směrem k vyšším hodnotám.

Z grafu Rozptýlení s kvantily též plyne, že v souboru měření se vyskytují vybočující hodnoty. Toto je na Obr. 1.8 zcela zřejmé, protože kvantilová funkce mimo obdélník E náhle roste.

V Kruhovém grafu (viz Obr.1.9) je zcela zřetelně vidět, že data jsou asymetrická a zešikmená směrem k vyšším hodnotám. Naměřená data též patrně nepocházejí z Normálního rozdělení

Závěr exploratorní analýzy:

Na základě exploratorních grafů je možno konstatovat, že naměřená data vykazují asymetrické rozdělení dat, zešikmené k vyšším hodnotám R_z a navíc patrně nepocházejícího z Normálního rozdělení.

Byly též nalezeny hodnoty, které by v případě datové symetrie mohly být chápány jako data vybočující. Protože se však jedná o rozdělení dat asymetrické není možno tato data vyloučit. V případném (a především neoprávněném) vyloučení těchto dat, by došlo ke ztrátě informace o tvaru rozdělení a navíc bychom uměle snížili datový rozptyl. Navíc zákazníka vždy zajímá nejvyšší hodnoty R_z , které slouží jako míra kvality povrchu reflektoru.

Použití odhadu aritmetického průměru jakožto odhadu střední hodnoty by bylo tedy zcela nesprávné (v důsledku asymetrie a vybočujících hodnot), kvalitnějším odhadem bude odhad mediánu a/nebo retransformovaného průměru.

Tab. 1. 2 Porovnání rozdělení naměřených dat R_z :

(1) LINEARITA V GRAFU KVANTIL-KVANTIL (Q-Q):				
Čís.	Rozdělení	Směrnice	Úsek	Korelační koeficient
0	Laplaceovo	6.0878E-01	2.5970E+00	9.3966E-01
1	Normální	8.0971E-01	2.5970E+00	9.3394E-01
2	Exponenciální	8.9232E-01	1.7265E+00	9.8169E-01
3	Rovnoměrné	2.6152E+00	1.2894E+00	9.0323E-01
4	Lognormální	5.0112E-01	1.8187E+00	9.6788E-01

Dle velikosti korelačního koeficientu (viz Tab.1.2) plyne, že nejvyšší hodnotu 0,9816 nabývá rozdělení Exponenciální, poté 0,96788 rozdělení Lognormální, 0,93966 rozdělení Laplaceovo, 0,93394 rozdělení Normální a nakonec rozdělení Rovnoměrné s korelačním koeficientem 0,90323

II Ověření předpokladů o výběrovém souboru dat hodnot R_z

Velikost výběrového souboru $n=30$

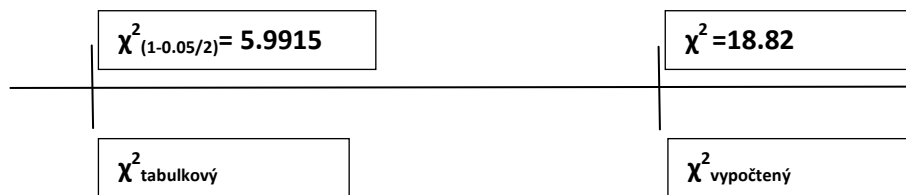
Hladina významnosti $\alpha=0.05$

Tab.2.1 Test normality

TEST NORMALITY:	
Tabulkový kvantil $\chi^2(1-\alpha,2)$: 5.9915E+00
χ^2 -statistika	: 1.8820E+01
Závěr: Předpoklad normality zamítnut	
Vypočtená hladina významnosti	: 8.1893E-05

Prvním krokem pro ověření předpokladů o výběrovém souboru dat je test normality. Jak plyne z Tab.2.1 hodnota tabulkového $\chi^2_{(1-0.05/2)}$ nabývá hodnoty 5.9915 oproti χ^2 vypočtené 18.82.

Při grafickém znázornění za účelem zlepšení vizualizace je možno vidět, že



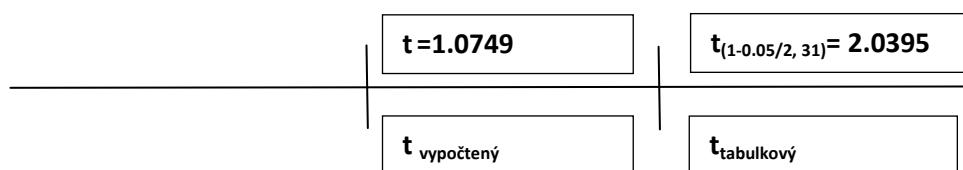
je nutné zamítnout hypotézu, že data mají Normálním rozdělení

Tab.2.2 Test nezávislosti naměřených dat

TEST NEZÁVISLOSTI:	
Tabulkový kvantil $t(1-\alpha/2, n+1)$: 2.0395E+00
Test autokorelace	: 1.0749E+00
Závěr: Předpoklad nezávislosti přijat	
Vypočtená hladina významnosti	: 1.4535E-01

Dalším krokem pro ověření předpokladů o výběrovém souboru dat je test nezávislosti. Jak plyne z Tab.2.2 hodnota tabulkového $t_{(1-0.05/2, 31)}$ nabývá hodnoty 2.0395 oproti t vypočtené 1.0749.

Při grafickém znázornění za účelem zlepšení vizualizace je možno vidět, že :



u analyzovaného výběru byla nezávislost prvků ve výběru prokázána.

Tab.2.3 Klasické odhady parametrů

KLASICKÉ ODHADY PARAMETRŮ:			
Průměr	: 2.5970E+00	Rozptyl	: 7.1003E-01
Sněrodatná odchylka:	8.4263E-01	Šikmost	: 1.3279E+00
Špičatost	: 4.7974E+00		

Dle Tab.2.3 opět můžeme specifikovat, že naměřená data jsou asymetrická (což mj. specifikuje i vysoká hodnota odhadu šikmosti (1,3279 um)).

Uvážíme-li ještě přítomnost vybočujících naměřených hodnot, můžeme konstatovat, že odhad aritmetického průměru 2,5970um nelze použít z důvodu jeho zkreslení.

Tab.2.4 Minimální velikost výběru a detekce vybočujících hodnot

MINIMÁLNÍ VELIKOST VÝBĚRU:			
pro 25% relativní chybu směrodatné odchylky	: n = 16		
pro 10% relativní chybu směrodatné odchylky	: n = 96		
pro 5% relativní chybu směrodatné odchylky	: n = 381		
DETEKCE ODLEHLÝCH BODŮ:			
Bod číslo 27 (horní)	: 5.2400E+00		
Počet odlehlých bodů	: 1		
Parametry s vynechanými odlehlými hodnotami:			
Průměr	: 2.5059E+00	Rozptyl	: 4.7731E-01
Sněrodatná odchylka:	6.9087E-01	Šikmost	: 8.8176E-01
Špičatost	: 3.5615E+00		

Jak plyne z Tab. 2.4 byl nalezena pouze jedna vybočující hodnota, a to $R_z = 5.24$ um, pořadové číslo měření 27.

Budeme-li uvažovat 25% relativní chybu směrodatné odchylky, bude minimální rozsah výběru $n=16$, pro 10% relativní chybu bude minimální velikost výběru $n = 96$ a pro 5% relativní chybu směrodatné odchylky bude $n = 381$.

Závěr z ověření předpokladů o výběrovém souboru dat hodnot R_z

Hodnota šikmosti $g_1=1,3279$ vede k předpokladu, že data pocházejí ze souboru asymetrického, což je v souladu s EDA (viz I).

Hodnota špičatosti $g_2=3,5615$ vede k předpokladu, že soubor dat nemá silnou špičatost, což je opět v souladu s EDA (viz I).

Dále se testem normality zjistilo, že je možné zamítnout předpoklad, že naměřená data pocházejí z Normálního rozdělení, a testem nezávislosti (pro $1-\alpha=0.95$) nebyla zamítnuta nezávislost dat (tedy data nepocházejí ze souboru v němž se vyskytuje autokorelace).

Byla nalezena pouze jedna vybočující hodnota (ovšem dle EDA viz I jsou vidět hodnoty 3 až 4), nicméně žádná z nich nemůže být ze souboru naměřených hodnot R_z vyloučena.

III Transformace dat

Z důvodu asymetrie naměřených dat a přítomnosti vybočující(ch) hodnot(y) je nutno provést Transformaci dat, která by měla vést ke stabilizaci rozptylu, zesymetričtění rozdělení a někdy i normalitě (1). Po provedené transformaci a nezbytných výpočtech je nutno provést opět transformaci inverzní.

Za účelem nalezení retransformovaného průměru, byly provedeny transformace:

- a) Transformace mocninná.
- b) Box-Coxova transformace.

ad a) Transformace mocninná

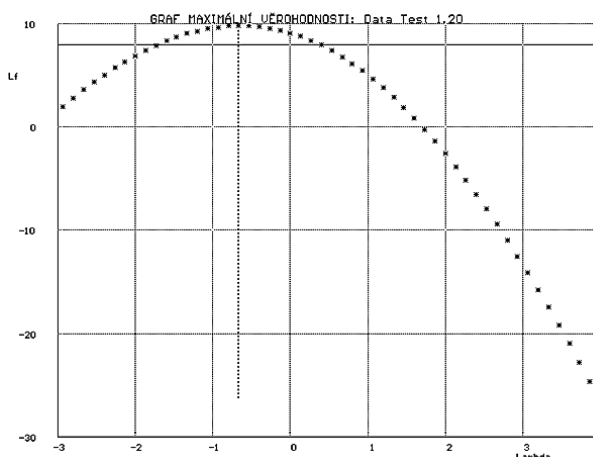
Tab.3.1 Prostá mocninná transformace a retransformované hodnoty

PROSTÁ MOCNINNÁ TRANSFORMACE:		
Optimální hodnoty mocniny pro vybraná kritéria:		
Optimální mocnina:-6.6667E-01	pro šikmost	: 3.9395E-03
Optimální mocnina:-1.3333E-01	pro špičatost	: 3.0310E+00
Optimální mocnina:-1.0667E+00	pro asymetrii	: 1.8199E-03
Optimální mocnina:-8.0000E-01	pro asymetrii, rob.	: 1.0513E-04
Optimální mocnina:-1.6000E+00	pro Hinkley-asymetrii:	1.6516E-03
Zvolená mocnina :-0.67		
Průměr	:	5.5526E-01
Rozptyl	:	1.1060E-02
Směrodatná odchylka	:	1.0517E-01
Šikmost	:	3.9395E-03
Špičatost	:	2.7721E+00
Opravený průměr	:	2.4169E+00

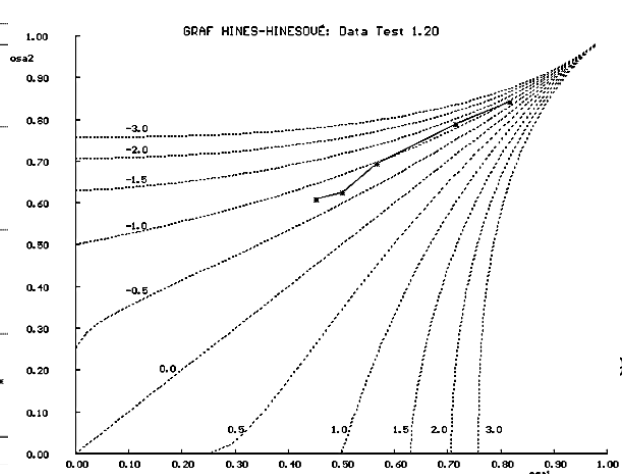
Výpočtem bylo zjištěno, že zvolená mocnina nabývá hodnoty o velikosti -0,67 (viz Tab.3.1). Obdobnou hodnoty je možné též získat (avšak pouze odhadem) pomocí grafu Hines Hinesová (viz Obr.3.2), kde získáme optimální odhad koeficientu λ .

Oprávněnost použití transformace je doloženo Grafem maximální věrohodnosti (viz Obr.3.1). Z grafu logaritmu věrohodnostní funkce je patrné, že interval parametru λ , $\langle \lambda_D, \lambda_U \rangle$, který byl sestaven na uvažování $1-\alpha=0.95$ nezahrnuje hodnotu 1. Tedy ze statistického hlediska je transformace naměřených dat hodnot R_z přínosem.

Obr.3.1. Graf maximální věrohodnosti



Obr.3.2 Graf Hines Hinesová



ad b) Transformace Box-Coxova

Tab.3.2 Box Coxova transformace a retransformované hodnoty

BOX-COXOVA TRANSFORMACE:		
Optimální hodnoty mocniny pro vybraná kritéria:		
Optimální mocnina:-6.6667E-01	pro šikmost	: 3.9395E-03
Optimální mocnina:-1.3333E-01	pro špičatost	: 3.0310E+00
Optimální mocnina:-1.0667E+00	pro asymetrii	: 1.8199E-03
Optimální mocnina:-8.0000E-01	pro asymetrii, rob.	: 1.0513E-04
Optimální mocnina:-1.6000E+00	pro Hinkley-asymetrii	: 1.0322E-03
Optimální mocnina:-6.6667E-01	pro věrohodnost	: 9.9109E+00
Zvolená mocnina :-0.67		
Průměr	:	6.6711E-01
Rozptyl	:	2.4885E-02
Sněrodatná odchylka	:	1.5775E-01
Šikmost	:	-3.9395E-03
Špičatost	:	2.7721E+00
Opravený průměr	:	2.4169E+00

Výpočtem bylo zjištěno, že zvolená mocnina nabývá hodnoty o velikosti -0,67 (viz Tab.3.2).

Závěr transformace dat

Tab.3.3 Opravené a klasické aritmetické průměry

Klasické odhady parametrů		Odhady parametrů po Prosté mocninné transformaci		Odhady parametrů po Box Coxově transformaci	
Průměr (um)	2.597	Optimální odhad koeficientu λ -0,67		Optimální odhad koeficientu λ -0,67	
		Opravený průměr (um)	2.4169	Opravený průměr (um)	2.4169

Oprávněnost použití transformace je doloženo Grafem maximální věrohodnosti (viz Obr.3.1). Z grafu logaritmu věrohodnostní funkce je patrné, že interval parametru λ , $\langle \lambda_D, \lambda_U \rangle$, který byl sestrojen na uvažováním $1-\alpha=0.95$, nezahrnuje hodnotu 1. Tedy ze statistického hlediska je transformace naměřených dat hodnot Rz přínosem.

Porovnání klasického odhadu aritmetického průměru a opravenými aritmetickými průměry je možné vysledovat v Tab. 3.3 včetně hodnot λ určených prostou a Box Coxovou transformací.

IV Analýza 1 výběru

Tab.4.1

V Ý S T U P	
PARAMETRY TVARU:	
Šikmost	: 1.3279E+00
Špičatost	: 4.7975E+00
KLASICKÉ ODHADY PARAMETRŮ :	
Průměr	: 2.5970E+00
Směr. odchylka	: 8.4263E-01
Rozptyl	: 7.1003E-01
95.0% spolehlivost:	
Spodní mez:	2.2824E+00 Horní mez: 2.9116E+00
ROBUSTNÍ ODHADY PARAMETRŮ :	
Medián	: 2.3750E+00
Směr. odchylka mediánu:	9.0121E-01
Rozptyl mediánu	: 8.1218E-01
Rozptyl (nepar.)	: 3.5636E-02
Směr. odchylka mediánu:	1.8878E-01
Rozptyl (Marritz)	: 5.7528E-02
Směr. odchylka mediánu:	2.3985E-01
95.0% spolehlivost:	
Spodní mez:	1.8845E+00 Horní mez: 2.8655E+00

Tab.4.2

Klasický odhad aritmetického průměru		Odhad aritmetického průměru po „Prosté mocninné transformaci“		Odhad aritmetického průměru po „Box Coxově transformaci“		Robusní odhad	
Průměr (um)	2.597	Opravený průměr	2.4169	Opravený průměr	2.4169	Odhad mediánu	2,375
Rz (um)		Rz (um)		Rz (um)		Rz (um)	

Závěr:

Naměřená data amplitudového parametru Rz (um) pocházejí z Laplaceova rozdělení (viz Tab.1.2), a proto nejlepší odhad střední hodnoty získáme transformací dat (mocninou či Box Coxovou) – viz Tab.4.2.

Můžeme konstatovat, že aplikacemi transformací (viz. Tab. 4,2), resp. robustním odhadem jsme získali přesnější hodnoty odhadů středních hodnot parametr jakosti povrchu Rz .

Statistická analýza malých výběrů dle Horna

Zadání:

Při testování creepu polymerního vzorku zhotoveného z ABS byl sledován parametr protažení v čase. Creepovací zařízení umožňovalo upnutí pouze čtyřech vzorků. Na základě těchto dat je třeba určit 95 % interval spolehlivosti střední hodnoty po 50 hodinách měření.

Data:

Tab. 1.1 Naměřená data

Čas (hod)	1.vzorek (mm)	2 vzorek (mm)	3 vzorek (mm)	4 vzorek (mm)
50	0,534	0,565	0,544	0,568

Tab.1.1 udává hodnoty protažení vzorků v mm za čas 50hod.

Užitý program: QC-Expert 2.9 CZ

Hornův postup:

1.Sestavení pořadové statistiky

Tab.1.2 Sestavená pořadová statistika

Čas[hod]	1. pořadová statistika	2. pořadová statistika	3. pořadová statistika	4. pořadová statistika
50	0,534 (mm)	0,544 (mm)	0,565(mm)	0,568(mm)

2.Hloubka pivotů H (pro $n = 4$)

$$H = \text{int}((n+1)/2+1)/2$$

$$\text{int} (4+1)/2+1)/2 = 1$$

3. Pivoty pro čas 50 hod

$$\text{Dolní pivot } x_D = x_{(H)} \quad x_D = x_{(1)} = 0,534 \text{ (mm)}$$

$$\text{Horní pivot } x_H = x_{(n+1-H)} \quad x_H = x_{(4+1-1)} = 0,568 \text{ (mm)}$$

4.Odhad parametru polohy pomocí pivotové polosumy

$$P_L = (x_D + x_H)/2 \quad P_L = (0,534 + 0,568)/2 = 0,551 \text{ (mm)}$$

5.Odhad parametru rozptýlení pomocí pivotového rozpětí

$$R_L = x_H - x_D \quad R_L = 0,568 - 0,534 = 0,034 \text{ (mm)}$$

6. 95% interval spolehlivosti střední hodnoty

$$P_L - R_L t_{L, 1-\alpha/2} (4) < \mu < P_L + R_L t_{L, 1-\alpha/2} (4)$$

$$0,551 - 0,034 \times 0,738 < \mu < 0,551 + 0,034 \times 0,738$$

$$0.5259 \text{ (mm)} < \text{střední hodnota} < 0.5761 \text{ (mm)}$$

Závěr:

Je možné konstatovat, že 95% interval spolehlivosti sestrojený Hornovou metodou pro střední hodnotu bude od $L_D = 0.5259$ (mm) do $L_H = 0.5761$ (mm).

Tab.1.3 Aplikace Softveru QCExpert - Analýza malých výběrů (Hornova metoda)

QC Expert Analýza malých výběrů (mm)	
Střední hodnota	0,551
Spodní mez (2.5%)	0,525908
Horní mez (97.5%)	0,576092
Pivotové rozpětí	3,39999999999999E-02

Tab. 1.3 udává výsledky určené softverem QC expert, které jsou ve shodě s „ručně“ vypočítanými parametry (viz odstavec 6.0).

Tab.1.4 Výpočet klasických parametrů softverem QCExpert

Klasické parametry (mm)	
Průměr	0,55275
Spodní mez	0,5265914335
Horní mez	0,5789085665
Směr. odchylka	0,01643928222
Šikmost	-0,1626053794
Špičatost	1,253961354

Tab.1.5 Výpočet robusních parametrů softverem QCExpert

Robustní parametry(mm)	
Medián	0,5545
IS spodní	0,5268966424
IS horní	0,5821033576
Medianová směr. odchylka	0,008673628705
Medianový rozptyl	7,523183492E-005

Tab.1.6 Porovnání klasické, robusní a Hornovy metody

	Poloha	Rozptýlení	95% interval spolehlivosti	
Klasická metoda	X	S	X_D	X_H
	mm	mm	Mm	mm
	0,553	0,0164	0,5266	0,5789
Robusní metoda	X_{0.5}	Medianová směr. odchylka	IS spodní	IS horní
	mm		Mm	mm
	0,555	0,0087	0,5269	0,5821
Hornova metoda	P_L	R_L	P_L - R_L t_{L, 1-α/2} (4)	P_L + R_L t_{L, 1-α/2} (4)
	0,551	0,034	0.5259	0.5761

Závěr:

S 95% statistickou jistotou můžeme tvrdit, že:

1. Střední hodnota μ protažení vzorku zhotoveného z polymerního materiálu ABS, zjištěná klasickými statistickými metodami, se nalézá v intervalu 0,5266 - 0,5789 (mm).
2. Střední hodnota M protažení vzorku zhotoveného z polymerního materiálu ABS, zjištěná použitím robusní statistiky, se nalézá v intervalu 0,5269- 0,5821 (mm).
3. Střední hodnota protažení vzorku zhotoveného z polymerního materiálu ABS, zjištěná Hornovou metodou, se nalézá v intervalu 0.5259 - 0.5761 (mm).

Test správnosti

Zadání:

Při hodnocení jakosti povrchu výrobku z ABS vyrobeného technologií vstřikování byl sledován amplitudový parametr drsnosti Ra dle ČSN EN ISO 4287. Velikost výběrového souboru $n = 60$.

Parametr Ra byl měřen kontaktním způsobem s využitím kontaktního profiloměru Taylor Hobson .

Zákazník vyžaduje, aby hodnota Ra byla dle výkresové dokumentace = 1,10 μm . Je jakost povrchu v souladu s požadavky zákazníka?

Data:

Tab. 1.1 Naměřené hodnoty amplitudového parametru drsnosti povrchu Ra (μm)

1,43	1	1,24	1,01	1,11	1,21	1,16	1,04	1,08	1,18
1,23	1,17	0,94	1,16	1,16	1,22	1,07	0,897	0,889	1,23
1,27	0,752	0,975	1,19	0,845	0,925	1,05	0,852	1,11	1,23
1,2	1,01	1,02	1,15	0,93	0,987	1	1,14	0,898	1,33
1,07	1,08	1,31	1,12	1,38	1,35	1,37	1,05	1,28	1,21
0,982	0,863	1,17	1,04	1,39	1,48	1,3	1,26	1,27	1,16

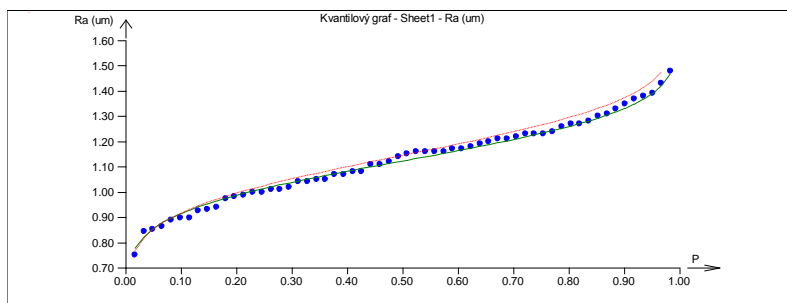
Užitý program: QC-Expert 2.9 CZ

Řešení:

I Průzkumová (explorativní) analýza dat

Odhalení stupně symetrie a špičatosti rozdělení

Obr.1.1 Kvantilový graf



V kvantilovém grafu – viz Obr 1.1 posuzujeme symetrii naměřených dat parametru Ra. Je možno konstatovat, že data se jeví symetrická, pouze v levé části Kvantilového grafu je jedna vybočující hodnota

Obr.1.2 Diagram rozptýlení a rozmítnutý diagram rozptýlení

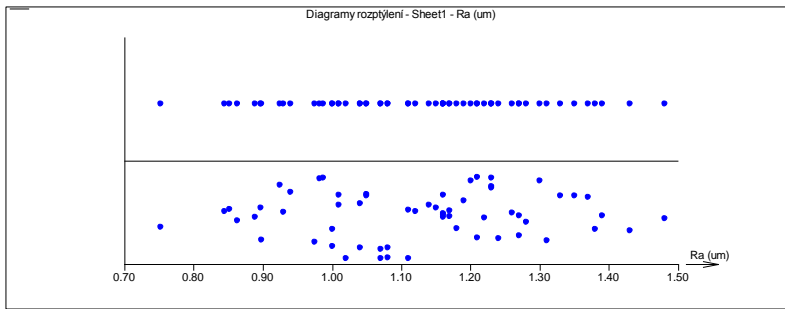
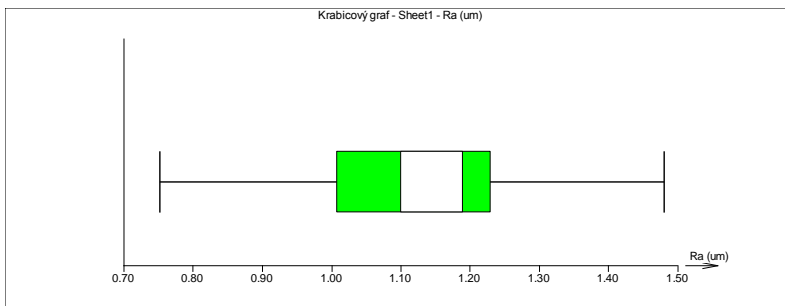


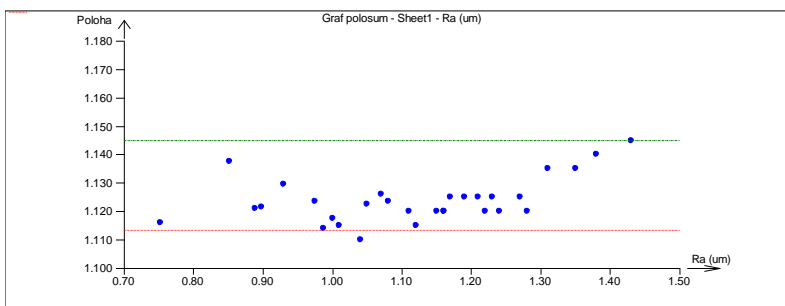
Diagram rozptýlení a rozmítnutý diagram rozptýlení (Obr. 1.2) detekuje symetrii naměřených dat parametru Ra. Je možno si též všimnout jedné vybočující hodnoty v dolní a dvou vybočujících hodnot v horní části diagramu.

Obr.1.3 Krabicový diagram s vousy



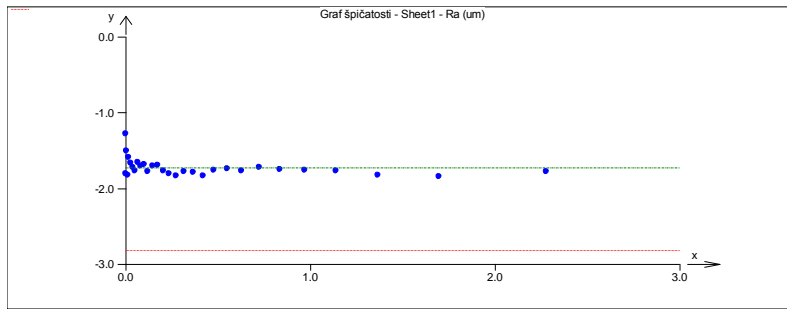
Krabicový diagram – viz Obr.1.3 (v literatuře uváděn jako s vousy, či anténami (boxplot with whiskers)) , detekuje symetrii naměřených dat parametru Ra. Dále tento diagram nespécifikuje žádné vybočující hodnoty parametru Ra.

Obr.1.4 Graf polosum



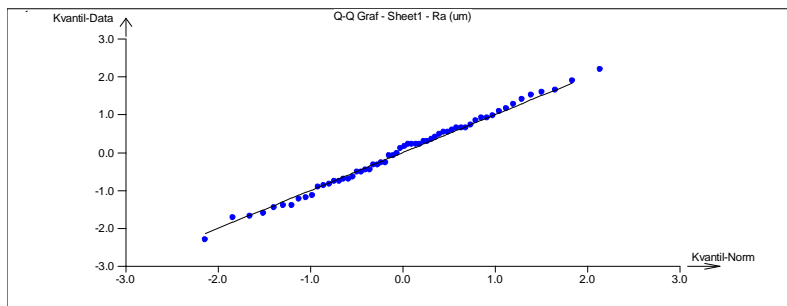
V Grafu polosum (Obr.1.4) vidíme, že zoom na ose y je relativně malý (1,11-1,145). Tvarově není výrazná odchylka od symetrického rozdělení, což právě potvrzuje i relativně malý zoom hodnot na ose y (tj. Poloha).

Obr.1.5 Graf špičatosti



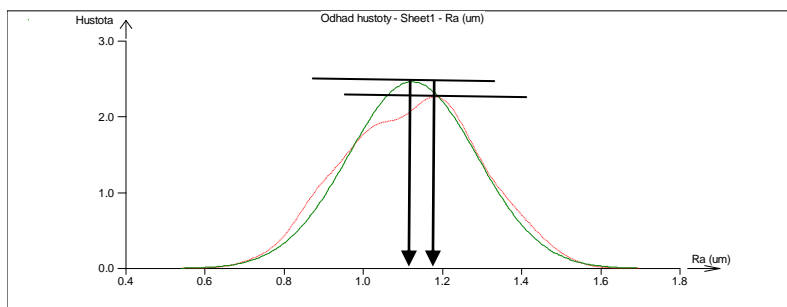
Dle Grafu špičatosti můžeme konstatovat, že většina bodů leží rovnoběžně s osou x. Jedná se tedy o symetricky rozdělená data.

Obr.1.6 Q-Q graf



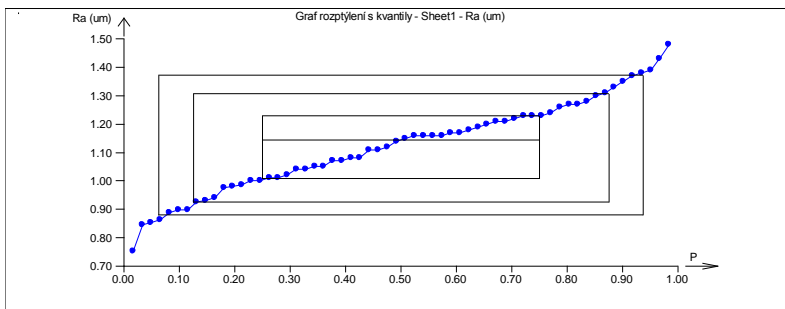
Při posouzení, zda data pocházejí ze souboru s normálním rozdělením využijeme QQ graf (viz Obr. 1.6). Z něho můžeme detekovat, že většina bodů (tedy naměřených hodnot parametru Ra) leží na teoretické přímce. Z tohoto můžeme usuzovat, že naměřená data pochází ze souboru, který má rozdělení normální.

Obr.1.7 Graf odhadu hustoty



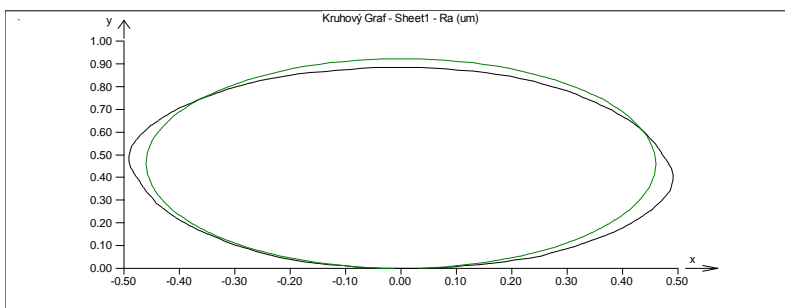
Při posuzování normality naměřených dat (tj. parametru Ra) můžeme dle Grafu odhadu jádrové hustoty (viz Obr. 1.7) specifikovat, že tvarové rozdělení není výrazně odlišné od rozdělení normálního. Průsečík kolmic spuštěných z vrcholů Gausovy křivky a křivky empirické na osu x nevykazuje též výrazné rozdíly.

Obr.1.8 Graf rozptýlení s kvantily



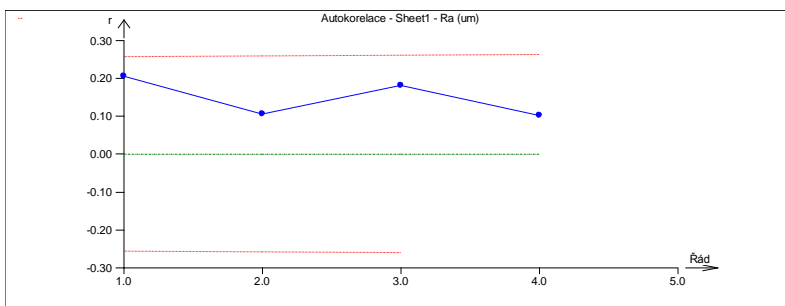
U Grafu rozptýlení s kvantily je vzdálenost dolních hran kvartilového obdelníku F, oktilového obdelníku E a sedecilového obdelníku D mírně kratší, než vzdálenost mezi horními hranami. Můžeme tedy usuzovat na mírné zešikmení směrem k nižším hodnotám.

Obr.1.9 Kruhový graf



V Kruhovém grafu (Obr. 1.9) je možné vidět mírné zešikmení směrem k nižším hodnotám. Dle tvaru grafů můžeme předpokládat, že naměřená data parametru Ra pocházejí z rozdělení normálního.

Obr.1.10 Graf autokorelace



Dle Grafu autokorelace můžeme usuzovat, že nedochází v naměřených datech k autokorelace od 1.ho do 4ho řádu.

II Ověření předpokladů o výběrovém souboru dat hodnot Ra

Velikost výběrového souboru $n = 60$

Hladina významnosti $1-\alpha = 0.95$

Tab.2.1 Test normality

Test normality :	
Název sloupce :	Ra (um)
Průměr :	1,12375
Rozptyl :	0,02623110593
Šikmost	-0,0367705611
Špičatost :	2,433813279
Normalita :	Přijata
Vypočtený :	0,06606128864
Teoretický :	5,991464547
Pravděpodobnost :	0,9675089105

Jak plyne z Tab. 2.1 hodnota tabulkového $\chi^2_{(1-0,05/2)}$ nabývá hodnoty 5,991464547 oproti vypočtené hodnotě $\chi^2 = 0,06606128864$. Protože 0,06606128864 (hodnota vypočtená) < 5,991464547 (hodnota tabulková) plyne závěr, že nezamítáme hypotézu, že naměřená data parametru Ra mají normální rozdělení.

Tab.2.2 Test nezávislosti naměřených dat

Znaménkový test	
Závěr :	Data jsou nezávislá

Dle znaménkového testu usuzujeme, že data jsou navzájem nezávislá.

Tab. 2.3 Test autokorelace naměřených dat:

Řád autokorelace 1	
Korelační koeficient :	0,2064108371
Pravděpodobnost :	0,05838318125
Závěr :	Nevýznamný
Řád autokorelace 2	
Korelační koeficient :	0,1061155237
Pravděpodobnost :	0,2139467545
Závěr :	Nevýznamný
Řád autokorelace 3	
Korelační koeficient :	0,1823317623
Pravděpodobnost :	0,08731316849
Závěr :	Nevýznamný
Řád autokorelace 4	
Korelační koeficient :	0,1032229655
Pravděpodobnost :	0,2245089251
Závěr :	Nevýznamný

Až do čtvrtého řádu autokorelace můžeme usoudit, že data tj. parametry Ra nejsou autokorelovaná.

Tab. 2.4 Test homogenity naměřených dat:

Vybočující body	
Homogenita :	Přijata
Počet vyb. bodů :	0
Spodní mez :	0,4963
Horní mez :	1,7337

Dle testu homogenity se v souboru parametru Ra (n=60) nevyskytují statisticky významně vychýlené hodnoty.

Tab. 2.5 Test významnosti trendu :

Směrnice :	0,001948680189
Významnost :	Nevýznamný
Pravděpodobnost :	0,9464587943

Dle tab. 2.5 můžeme specifikovat, že trend naměřených dat parametru Ra je nevýznamný.

Závěr z ověření předpokladů o výběrovém souboru dat hodnot Ra

Na základě ověření předpokladů o výběrovém souboru amplitudového parametru Ra o n=60 můžeme konstatovat, že se jedná o soubor dat vykazující symetrii, pocházející z Normálního rozdělení, vykazující nezávislost, splňující podmínky homogenity a nevykazující autokorelaci (do 4.ho řádu).

III Analýza výběrového souboru parametru Ra (um)

Tab. 3.1 Srovnání klasických a robustních parametrů

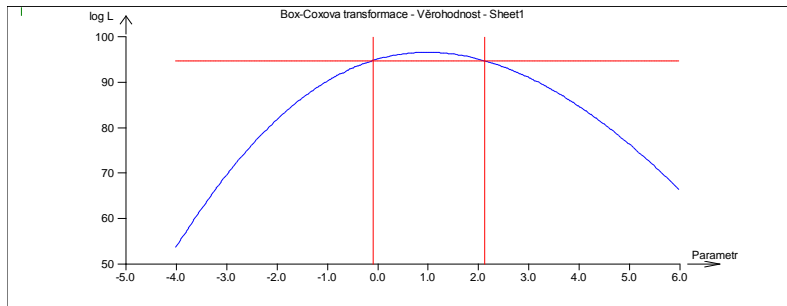
Klasické parametry Ra (um)		Robustní parametry Ra (um)	
Průměr :	1,12375	Medián :	1,145
Spodní mez :	1,081911244	IS spodní :	1,073534567
Horní mez :	1,165588756	IS horní :	1,216465433
Rozptyl :	0,02623110593	Medianová směr. odchylka :	0,03571494173
Směr. odchylka :	0,1619601986	Medianový rozptyl :	0,001275557063
Šikmost	-0,0367705611		
Odchylka od 0 :	Nevýznamná		
Špičatost :	2,433813279		
Odchylka od 3 :	Nevýznamná		
Polosuma	1,116		
Modus :	1,186106557		

Závěr:

Z Tab.3.1 je možné konstatovat, že 95% interval spolehlivosti pro μ bude od $L_D = 1,08$ (um) do $L_H = 1,17$ (um). Dále 95% interval spolehlivosti pro M bude od $L_D = 1,07$ (um) do $L_H = 1,22$ (um). A dále konstatujeme, že odchylka od šikmosti i špičatosti je nevýznamná.

IV Transformace dat (provedena pouze z cvičných důvodů)

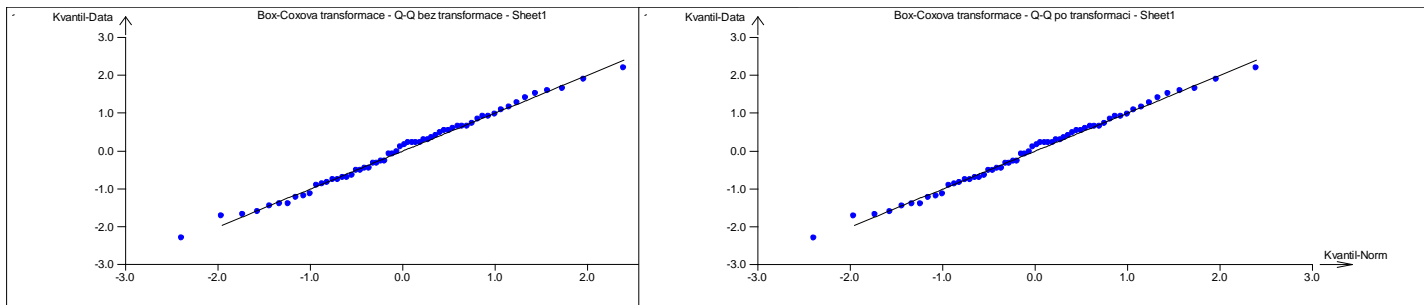
Obr. 4.1 Graf maximální věrohodnosti



Jak plyne z grafu maximální věrohodnosti (viz. Obr.4.1), přesněji řečeno z grafu logaritmu věrohodnostní funkce (s uvažováním $1-\alpha = 0,95$), v červeném intervalu leží hodnota 1. To značí, že použití Box Coxovy transformace ze statistického hlediska není přínosem.

Obdobný předpoklad můžeme vyslovit i z Q-Q grafů před a po transformaci – viz obr 4.1.

Obr.4.1 Q-Q Grafy před a po Box Coxově transformaci dat parametru Ra



Tab.4.1. Parametry získané transformací

Box-Coxova transformace dat		Odhad ar.průměru parametru Ra (um)	Odhad mediánu parametru Ra (um)
Optimální parametr :	0,9918178558		
Dolní mez parametru :	-0,08858417617		
Horní mez parametru :	2,122471088		
Věrohodnost bez transformace :	96,56627475		
Věrohodnost s transformací :	96,56637429		

Oprávněnost transformace :	Ne	Odhad ar.průměru parametru Ra (um)	Odhad mediánu parametru Ra (um)
Pravděpodobnost :	3,26600358095797 %		
Zvolený parametr :	0,9918178558		
Věrohodnost :	96,56637429		
Opravený průměr :	1,123617125	1,12375	1,145
LCL :	0,6391952416		
UCL :	1,610589175		
LWL :	0,9617720905		
UWL :	1,285729074		

Je možné konstatovat, že rozdíl věrohodnosti bez transformace a s transformací se liší až na čtvrtém desetinném místě, a tedy předpokládáme nevýznamnost tohoto rozdílu.

Jak plyne z výsledků (viz Tab.4.1), retransformovaný aritmetický průměr parametru Ra, $X_{Ra} = 1,123617125(\mu\text{m})$ a aritmetický průměr $X = 1,12375(\mu\text{m})$ se liší až na místě čtvrtém za desetinnou tečkou (viz.Tab.4.1).

Závěr:

Z Tab.3.1 je možné dále konstatovat, že 95% interval spolehlivosti pro μ bude od $L_D = 1,08(\mu\text{m})$ do $L_H = 1,17(\mu\text{m})$. Dále 95% interval spolehlivosti pro M bude od $L_D = 1,07(\mu\text{m})$ do $L_H = 1,22(\mu\text{m})$.

Tyto intervaly pokrývají zákazníkem požadovanou hodnotu Ra = 1,1 μm , proto můžeme se statistickou jistotou ($1 - \alpha = 0,95$) tvrdit, že požadavky zákazníka byly splněny.

Test shodnosti

Zadání:

Pro hodnocení jakosti povrchů vzniklých po aplikaci technologie jemné broušení byla zvolena technika otisku pomocí otiskovací hmoty o komerčním názvu Dental Flex, využívanou běžně ve stomatologii.

Po provedení otisku byla hodnocena jakost povrchu vzorku a i otisku pomocí amplitudového parametru Ra v μm , dle ISO 4287 a ISO 4288.

Pomocí teorie hypotéz chceme zjistit, zda stomatologická hmota Dental Flex je využitelná pro hodnocení jakosti povrchů po aplikaci technologie jemného broušení.

Data

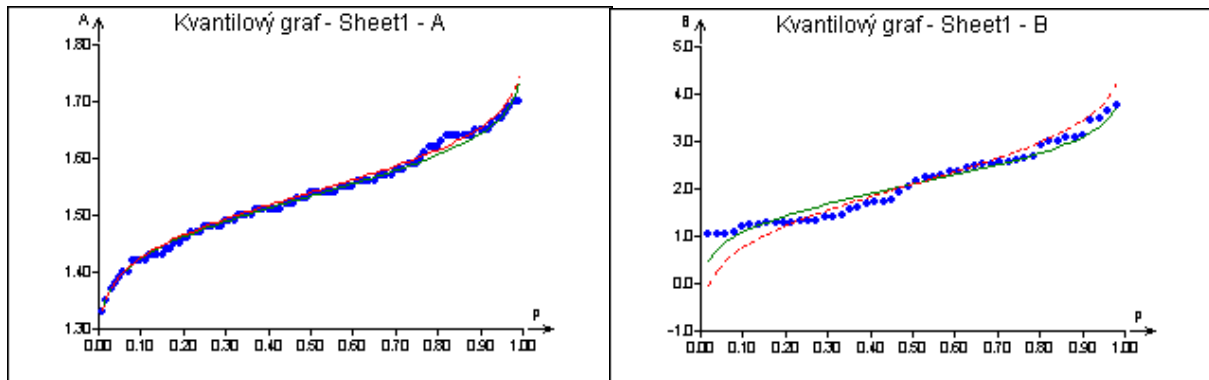
1,58	1,51	1,48	1,46	1,52	1,54	1,43	1,39	1,4	1,42	1,49	1,52	1,59	1,65	1,62
1,5	1,55	1,45	1,48	1,35	1,46	1,49	1,45	1,53	1,54	1,53	1,51	1,67	1,62	1,64
1,64	1,55	1,7	1,65	1,54	1,49	1,51	1,6	1,66	1,54	1,59	1,65	1,58	1,55	1,5
1,64	1,63	1,68	1,61	1,64	1,64	1,48	1,43	1,47	1,42	1,43	1,42	1,48	1,55	1,56
1,57	1,56	1,65	1,64	1,4	1,42	1,37	1,33	1,38	1,44	1,57	1,57	1,56	1,58	1,54
1,7	1,69	1,62	1,64	1,54	1,59	1,67	1,56	1,51	1,57	1,54	1,47	1,48	1,44	1,56
1,53	1,5	1,51	1,43	1,47	1,51	1,52	1,5	1,51	Vzorek Ra (μm)					
2,25	1,05	2,35	2,57	3,09	1,58	1,74	3,08	1,73	3,43	2,04	3,63	2,58	1,34	2,25
1,69	1,43	2,94	1,29	1,31	3,02	1,71	2,57	1,08	2,65	2,45	1,22	2,16	1,27	3,11
1,29	1,57	1,05	2,69	2,52	2,53	1,33	1,92	1,4	1,29	3,74	1,39	1,03	2,36	1,23
3	2,3	3,48	2,47	1,2	Otisk Ra (μm)									

Užitý program: QC-Expert 2.9 CZ, Adstat125-Student, MINITAB 14

Řešení

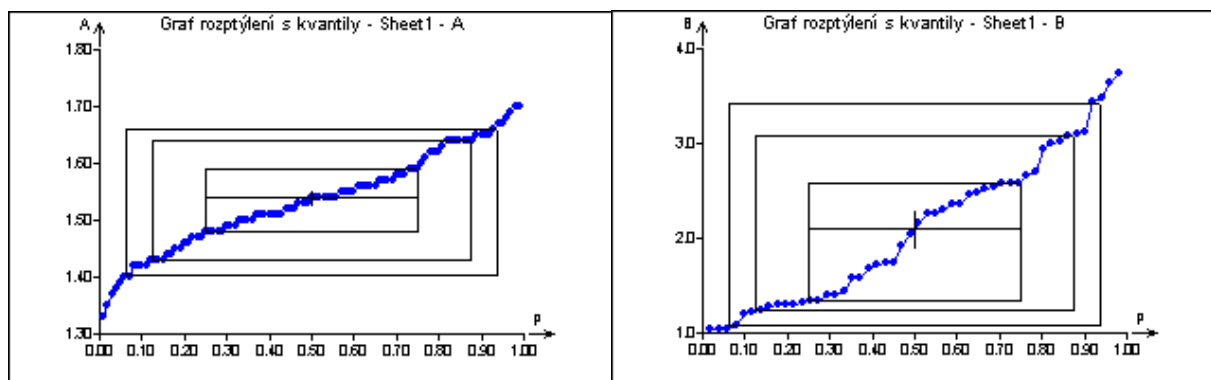
Průzkumová (explorativní) analýza dat

Obr.1.1 Kvantilové grafy vzorku a otisku



Při hodnocení kvantilových grafů vzorku (označen jako A) a jeho otisku (označen jako B) je vidět diametrální rozdíl. Vzorek (A) nevykazuje odlehle body a ukazuje na dobrou datovou symetrii parametru R_a , otisk (B) se jeví jako méně symetrický, v porovnání se vzorkem (A) (viz Obr.1.1).

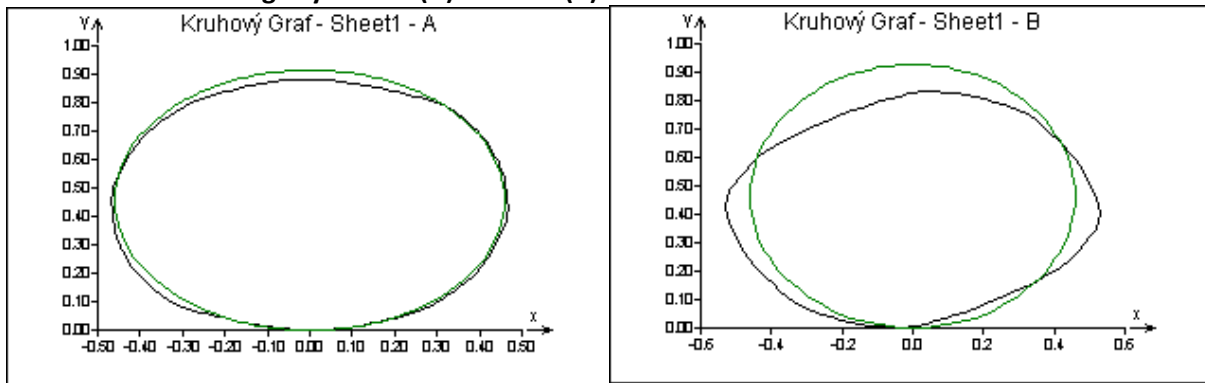
Obr.1.2 Graf rozptýlení s kvantily vzorku (A) a otisku (B)



U grafů rozptýlení s kvantily má kvantilová funkce vzorku (A) (viz Obr. 1.2A) přibližně sigmoidální tvar, vzdálenosti dolních a horních hran kvartilového, oktilového a sedecilového obdélníku jsou přibližně stejné (graf tedy vykazuje symetrii naměřených dat parametru R_a).

Oproti tomu u otisku (B) nemá kvantilová funkce (viz Obr. 1.2 B) sigmoidální tvar, vzdálenosti dolních hran kvartilového, oktilového a sedecilového obdélníku jsou kratší, než vzdálenosti mezi jejich horními hranami. Jedná se tedy o zešikmení směrem k vyšším hodnotám parametru R_a .

Obr.1.3 Kruhové grafy vzorku (A) a otisku (B)



Kruhové grafy (viz obr 1.3) potvrzují výsledky získané pomocí Grafů rozptýlení s kvantily. V případě vzorku (A) se jedná prakticky o data s minimální asymetrií, naopak u otisku (B) je asymetrie dat zřetelná.

Závěr explorativní analýzy:

Je možné konstatovat, že na základě tří typů grafů vytvořených pro vzorek (A) a jeho otisk (B), data otisku vykazují asymetrické rozdělení naměřeného parametru Ra, oproti vzorku (A), kde se asymetrie výrazněji neprojevila.

Dle použitých grafů je možné též konstatovat, že vybočující hodnoty nebyly nalezeny u vzorku (A), u otisku (B) je posouzení problematické. Ve spojení s kvantilovým grafem by se dalo usoudit na tři vybočující hodnoty, které by mohly způsobovat asymetrii. Zde je nutný podrobnější rozbor – viz ověření předpokladů.

Otázka normality dat je poněkud sporná. Z grafů se dá usuzovat, že data vzorku (A) mají rozdělení normální, zatímco data otisku (B) nikoliv.

Ověření předpokladů:

Test normality :

Byl proveden pomocí softwaru QCExpert s následujícími výsledky – viz. Tab.2.1

Tab.2.1 Test normality dat Ra vzorku (A) a otisku (B)

Název sloupce :	Ra_A_Vzorek (um)	Ra_B_Otisk (um)
Průměr :	1,533737374	2,088
Rozptyl :	0,007305277262	0,6118734694
Šikmost	-0,06391427364	0,3588889384
Špičatost :	2,369761811	1,976859363

Normalita :	Přijata	Přijata
Vypočtený :	0,1330235604	1,531518419
Teoretický :	5,991464547	5,991464547
Pravděpodobnost :	0,9356518936	0,4649807792

Při hodnocení výsledků z Tab.2.1 plyne, že pro data Ra ve sloupcích označených identifikátory Ra_A_Vzorek (um), Ra_B_Otisk (um) je Normalita přijata ($1-\alpha = 0.95$).

Hodnocení s využitím Tab.2.1

$$\text{Ra_A_Vzorek: } \chi^2_{(1-0.05/2)} = 5,991464547 > \chi^2 = 0,1330235604$$

Protože hodnota teoretická $5,991464547 > 0,1330235604$ není možné ($1-\alpha$) zamítnout normalitu dat parametru Ra vzorku (A). Tomu i odpovídá pravděpodobnost $0,9356518936 > 0,050$.

$$\text{Ra_B_Otisk: } \chi^2_{(1-0.05/2)} = 5,991464547 > \chi^2 = 1,531518419$$

Protože hodnota teoretická $5,991464547 > 1,531518419$ není možné ($1-\alpha$) zamítnout normalitu dat parametru Ra otisku (B). Tomu i odpovídá pravděpodobnost $0,4649807792 > 0,050$.

Dále podrobíte testování dat otisku (B) na stejné $1-\alpha = 0,95$ softverem ADStat.

Při použití programu ADStat obdržíme pro tatáž data (tj. B_Otisk) výsledek:

Tab.2.2 Test normality a Nezávislosti Otisku (B)

TEST NORMALITY:

```
Tabulkový kvantil Chi^2(1-alfa,2) : 5.9915E+00
Chi^2-statistika : 3.5054E+00
Závěr: Předpoklad normality přijat
Vypočtená hladina významnosti : 1.7331E-01
```

TEST NEZÁVISLOSTI:

```
Tabulkový kvantil t(1-alfa/2,n+1) : 2.0076E+00
Test autokorelace : 1.3030E+00
Závěr: Předpoklad nezávislosti přijat
Vypočtená hladina významnosti : 9.9217E-02
```

I zde dochází k přijetí normality a předpokladu nezávislosti.

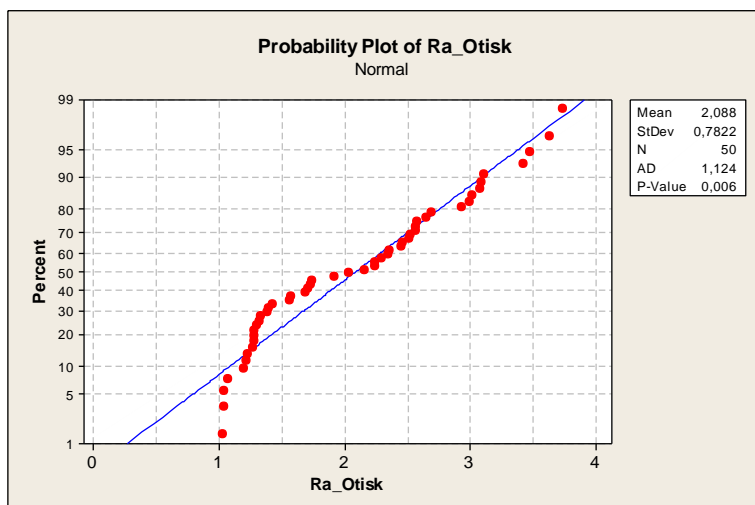
Dále využijí pro testování normality softwaru Minitab ver.14.0

Tatáž data (tj. pro otisk (B))podrobíme testu Anderson Darlingovu (pro stejnou hodnotu $1-\alpha = 0.95$).

Jak plyne z Obr.2.1 zde je nezbytné Normalitu zamítnout, protože P- hodnota (dle amerického testování) nabývá hodnoty pouze 0,006 což je numericky o řád méně než pro $\alpha = 0.05$.

$$\text{Tedy } P = 0,006 < \alpha = 0,050$$

Obr.2.1 Anderson Darlinguv test normality pro otisk (B)

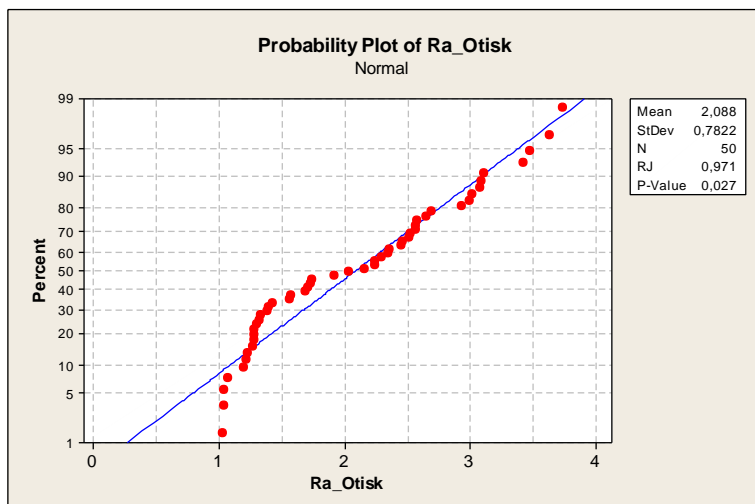


Pokračujme ještě dalším testem, a to Ryan Joinerův test, který však pro četnost $n = 50$ není příliš vhodný.

Jak plyne z Obr.č.2.2 i zde je nutno zamítnout hypotézu o Normalitě dat.

Tedy P- hodnota = $0.026 < \alpha = 0,050$.

Obr.2.2 Ryan Joinerův test normality pro otisk (B)



Při hodnocení sloupce B_Otisk, s využitím třech statistických programů (ADStat, QCExpert a Minitab) a třech testů (Ryan Joinerův, Anderson Darlingův a kombinovaný test šikmosti a špičatosti) a dále metodiky EDA, můžeme konstatovat, že Normalitu dat ve sloupci B_Otisk je možno zamítnout u prvních dvou testů (tj. Ryan Joinerův a Anderson Darlingův, avšak zamítnutí není možné při aplikaci kombinovaného testu šikmosti a špičatosti)

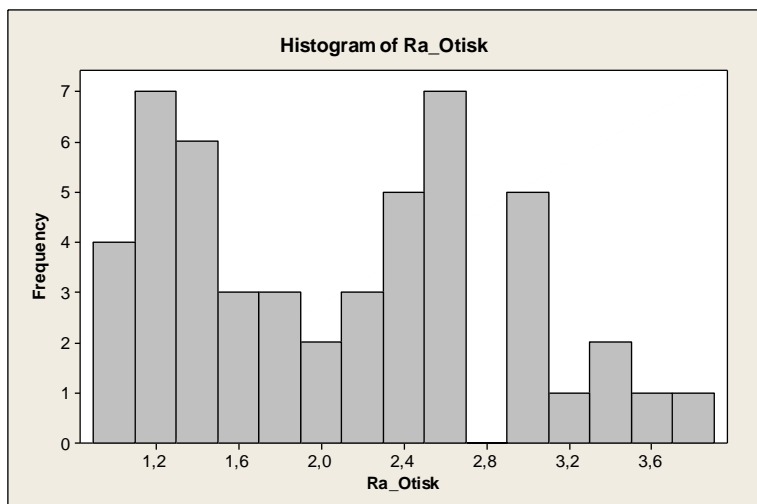
Tab.2.3 Porovnání rozdělení naměřených dat parametru Ra otisku (B)

Čís.	Rozdělení	Směrnice	Úsek	Korelační koeficient
0	Laplaceovo	5.4627E-01	2.0880E+00	9.3146E-01
1	Normální	7.7477E-01	2.0880E+00	9.7143E-01
2	Exponenciální	7.7273E-01	1.3276E+00	9.3636E-01
3	Rovnoměrné	2.6534E+00	7.6128E-01	9.8406E-01
4	Lognormální	3.8859E-01	1.4738E+00	8.5895E-01
5	Gumbe lovo	5.7711E-01	2.4139E+00	9.1167E-01

Celou situaci dále komplikuje i Tab.2.3, kde dle odhadu korelačního koeficientu můžeme stanovit, že nejvyšší hodnotu (korelačního koeficientu) nabývá rozdělení Rovnoměrné (0,984), za ním následuje Normální (0,971).

Navíc při posouzení rozdělení dat sloupce B_Otisk, pomocí histogramu vzniká podezření na bimodalitu dat (viz Obr.2.3).

Obr.2.3 Histogram naměřených dat Ra pro otisk (B)



Homogenita naměřených dat

Tab. 2.4 Vybočující hodnoty nalezeny softverem QC Expert

Název sloupce :	Ra_A_Vzorek	Ra_B_Otisk
Homogenita :	Přijata	Přijata
Počet vybočujících hodnoty :	0	0
Spodní mez :	1,204363636	-1,43428
Horní mez :	1,855636364	5,31428

Jak plyne z Tab.2.4, vybočující hodnoty nebyly nalezeny.

Test nezávislosti naměřených dat

Tab.2.5 Test nezávislosti pro Ra vzorek (A) pomocí softveru ADStat Student

TEST NEZÁVISLOSTI :

Tabulkový kvantil $t(1-\alpha/2, n+1)$: 1.9840E+00
Test autokorelace : 1.0117E+00
Závěr: Předpoklad nezávislosti přijat
Vypočtená hladina významnosti : 1.5705E-01

$$t_{\text{tabulkový}} = 1,9840 > t_{\text{vypočítaný}} = 1,0117$$

Tab.2.6 Test nezávislosti pro Ra otisk (B)

TEST NEZÁVISLOSTI :

Tabulkový kvantil $t(1-\alpha/2, n+1)$: 2.0076E+00
Test autokorelace : 1.3030E+00
Závěr: Předpoklad nezávislosti přijat
Vypočtená hladina významnosti : 9.9217E-02

$$t_{\text{tabulkový}} = 2,0076 > t_{\text{vypočítaný}} = 1,303$$

Závěr:

Jak plyne z Tab. 2.5 a 2.6 a s využitím výše uvedených softverů je možné konstatovat, že data Ra vzorku (A) a otisku (B) nejsou závislá ($1-\alpha = 0,95$)

Odhady klasických parametrů:

Posledním přípravným krokem je určení klasických odhadů parametrů viz Tab.2.7 a parametrů robusních viz Tab.2.8.

Tab. 2.7 Odhady Klasických parametrů :

Název sloupce :	Ra_Vzorek_A (um)	Ra_Otisk_B (um)
Počet platných dat :	99	50
Průměr :	1,533737374	2,088
Spodní mez :	1,5166905	1,865694565
Horní mez :	1,550784247	2,310305435
Rozptyl :	0,007305277262	0,6118734694
Směr. odchylka :	0,08547091472	0,782223414

Šikmost	-0,06391427364	0,3588889384
Odchylka od 0 :	Nevýznamná	Nevýznamná
Špičatost :	2,369761811	1,976859363
Odchylka od 3 :	Nevýznamná	Nevýznamná

Tab 2.8 Robustní parametry :

Název sloupce :	Ra_Vzorek_A (um)	Ra_Otisk_B (um)
Medián :	1,54	2,1
IS spodní :	1,51974996	1,49506573
IS horní :	1,56025004	2,70493427
Medianová směr. odchylka :	0,01020426907	0,3010259374
Medianový rozptyl :	0,0001041271072	0,090616615

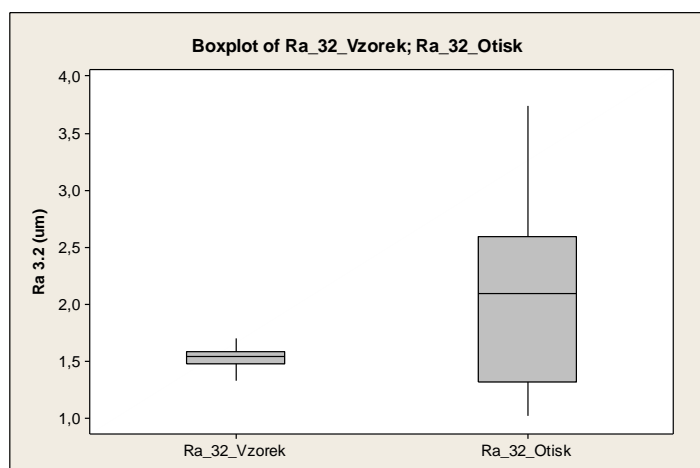
Testování dat parametru Ra vzorku (A) a otisku (B) pomocí statistických hypotéz

Test shody rozptylů ($\alpha = 0,05$)

Před vlastním testem rozptylů, provedeme zobrazení naměřených dat parametru Ra_Vzorek_A a Ra_Otisk_B pomocí krabicových diagramů – viz Obr.3.1. Oba grafy mají shodné měřítko a je vidět (obdobně jako v metodice EDA), že rozptyly naměřených hodnot Ra se liší (tím je míněno IQR, R, s apod.).

Toto je však nutné dokázat exaktně, tedy s využitím F-testu spadajícího do statistické teorie hypotéz.

Obr.3.1 Krabicový diagram (Box Plot) pro data vzorku (A) a otisku (B)



I Varianta:

F testy jsou obecně „citlivé“ na dodržení podmínek normality. V tomto případě budeme předpokládat výsledky z pasáže Ověření předpokladů. Tedy na základě kombinovaného testu šikmosti a špičatosti nezamítáme normalitu dat jak u vzorku (A), tak i otisku (B).

Stanovíme nulovou a alternativní hypotézu:

$$H_0 : \sigma^2_{Ra_A_Vzorek} = \sigma^2_{Ra_A_Otisk}$$

$$H_a : \sigma^2_{Ra_A_Vzorek} \neq \sigma^2_{Ra_A_Otisk}$$

Dále zvolíme konfidenční úroveň:

$$\text{Pro } 1-\alpha = 0.95$$

Tab.3.1 parametry F testu

Fisher–Snedecor F–test:

Počet stupňů volnosti Df1	:	49
Df2	:	98
Tabulkový kvantil $F(1-\alpha/2, Df1, Df2)$:	1.5990E+00
F–statistika	:	8.3758E+01
Závěr: Rozptyly se považují za rozdílné, H_0 zamítnuta		
Uypočtená hladina významnosti	:	0.000

Statistický závěr:

Z Tab.3.1 plyne, že zamítáme H_0 o rovnosti rozptylů Ra ve sloupcích označených jako Vzorek_A_Ra a Otisk_B-Ra. Dle amerického testování je pravděpodobnost $p = 0,000 < \alpha = 0,050$.

Stejný závěr (tedy zamítnutí H_0) udává i Tabulkový kvantil $F = 1.5990 < F \text{ statistika} = 83.758$

Praktický závěr:

Je možno konstatovat, že rozptyly dat parametru Ra vzorku (A) a jeho otisku (B) jsou rozdílné (což plyne mj. i z Obr.3.1).

II Varianta:

V případě, že budeme předpokládat výsledky z pasáže Ověření předpokladů, potom na základě Ryan Joinerova a Anderson Darlingova testu normalitu dat u otisku (B) zamítáme. Dále bereme do úvahy i Tab.2.3.

V tomto případě provedeme korigovaný test shody rozptylů viz Tab.3.2

3.2 Korigovaný F test shody rozptylů

Korigovaný F-test :
Počet stupňů volnosti Df1 : 42
Df2 : 21
Tabulkový kvantil $F(1-\alpha/2, Df1, Df2)$: 2.2374E+00
F-statistika : 8.3758E+01
Závěr: Rozptyly se považují za rozdílné, H_0 zamítnuta
Vypočtená hladina významnosti : 0.000

Statistický závěr:

Z Tab.3.2 plyne, že zamítáme H_0 o rovnosti rozptylů σ_A . Dle amerického testování je pravděpodobnost $p = 0,000 < \alpha = 0,050$.

Stejný závěr (tedy zamítnutí H_0) udává i Tabulkový kvantil $F = 2.2374 < F \text{ statistika} = 83.758$

Praktický závěr:

Je možno konstatovat, že rozptyly dat parametru σ_A vzorku (A) a jeho otisku (B) jsou rozdílné .

Celkový závěr:

Je možné konstatovat, že vzorek (A) a jeho otisk (B) mají rozdílné rozptyly ($1-\alpha=0,95$) jak při řešení **variantou I.** (uvažují normální rozdělení vzorku (A) i otisku (B)), tak i **variantou II** (uvažují normální rozdělení vzorku (A) i jiné než normální u otisku (B)).

Test shody průměrů pro ROZDÍLNÉ rozptyly

I.Varianta – předpokládám, že data vzorku (A) a otisku (B) mají normální rozdělení a rozdílné rozptyly

Stanovíme nulovou a alternativní hypotézu:

$$H_0 : \mu_{Ra_A_Vzorek} = \mu_{Ra_A_Otisk}$$

$$H_a : \mu_{Ra_A_Vzorek} \neq \mu_{Ra_A_Otisk}$$

Dále zvolíme konfidenční úroveň:

$$\text{Pro } 1-\alpha = 0.95$$

Tab. 3.3 Parametry t-testu

t-test (pro různé rozptyly)

Počet stupňů volnosti Df1 : 51
Tabulkový kvantil $t(1-\alpha/2, Df1)$: 2.0076E+00
t-statistika : 4.9953E+00
Závěr: Průměry se považují za rozdílné, H_0 zamítnuta
Vypočtená hladina významnosti : 0.000

Statistický závěr:

Z Tab.3.3 plyne, že zamítáme H_0 o rovnosti středních hodnot parametru Ra (na hladině významnosti $\alpha = 0.050$). Dle amerického testování je pravděpodobnost $p = 0,000 < \alpha = 0,050$.

Stejný závěr (tedy zamítnutí H_0) udává i Tabulkový kvantil $t = 2.0076 < t_{\text{statistika}} = 4.9953$

Praktický závěr: Je možno konstatovat, že střední hodnoty dat parametru Ra vzorku (A) a jeho otisku (B) jsou rozdílné.

II. Varianta – předpokládám, že data vzorku (A) mají normální rozdělení, data otisku (B) rozdělení normální nemají. Dále předpokládám rozdílné rozptyly.

Stanovíme nulovou a alternativní hypotézu:

$$H_0 : \mu_{Ra_A_Vzorek} = \mu_{Ra_A_Otisk}$$

$$H_a : \mu_{Ra_A_Vzorek} \neq \mu_{Ra_A_Otisk}$$

Dále zvolíme konfidenční úroveň:

$$\text{Pro } 1-\alpha = 0.95$$

Tab. 3.4 Parametry modifikovaného t-testu

t-test(modifikovaná šikmost)

U-star	: 4.5685E+00
B-X charakteristika	: 2.0452E-05
B-Y charakteristika	: 4.2676E-01
Počet stupňů volnosti Df1	: 51
Tabulkový kvantil t(1-alfa/2,Df1)	: 2.0076E+00
t-statistika	: 4.5685E+00
Závěr: Průměry se považují za rozdílné, H_0 zamítnuta	
Vypočtená hladina významnosti	: 0.000

Statistický závěr:

Z Tab.3.4 plyne, že zamítáme H_0 o rovnosti středních hodnot parametru Ra (na hladině významnosti $\alpha = 0.050$). Dle amerického testování je pravděpodobnost $p = 0,000 < \alpha = 0,050$.

Stejný závěr (tedy zamítnutí H_0) udává i Tabulkový kvantil $t = 2.0076 < t_{\text{statistika}} = 4.5685$

Praktický závěr: Je možno konstatovat, že střední hodnoty dat parametru Ra vzorku (A) a jeho otisku (B) jsou rozdílné.

Celkový závěr:

Je možné konstatovat, že vzorek (A) a jeho otisk (B) mají rozdílné střední hodnoty se statistickou jistotou danou $1-\alpha = 0,95$ jak při řešení **variantou I**. (uvažují normální rozdělení vzorku (A) i otisku (B)), tak i **variantou II** (uvažují normální rozdělení vzorku (A) i jiné než normální u otisku (B)).

Závěr tedy zní, že dentální otiskovací hmota o komerčním názvu Dental Flex není využitelná pro hodnocení jakosti jemných povrchů vzniklých aplikací technologie broušení.

Párový test

Zadání:

Na formu pro vstřikování polymerních materiálů byl nanesen dezén technikou chemického leptání. Forma poté byla v provozu a vykonala 500 000 cyklů, čímž došlo k jejímu opotřebení. Kvalita dezénu byla hodnocena amplitudovým parametrem Rz (um). Je možné specifikovat, že po 500 000 cyklech je dezén opotřeben statisticky významně?

Data

Dezén nové formy Rz (um)	Dezén formy po 500 000 cyklech Rz (um)	d(i) Rz (um)
2,64	2,48	0,16
2,66	2,38	0,28
2,76	2,41	0,35
2,58	2,40	0,18
2,69	2,37	0,32
2,70	2,37	0,33
2,70	2,51	0,19
2,66	2,35	0,31
2,60	2,51	0,09
2,53	2,47	0,06

Užitý program: QC-Expert 2.9 CZ

Řešení:

Stanovíme nulovou a alternativní hypotézu a taktéž hladinu významnosti α .

$$H_0: \bar{d} = 0$$

$$H_a: \bar{d} \neq 0$$

$$1 - \alpha = 0,95$$

Tab. 1.1 Klasické parametry:

Klasické parametry		
	Dezén nové formy Rz (um)	Dezén formy po 500 000 cyklech Rz (um)
Průměr :	2,652	2,425
Spodní mez :	2,603858612	2,38099925
Horní mez :	2,700141388	2,46900075
Rozptyl :	0,004528888889	0,003783333333
Směr. odchylka :	0,06729701991	0,06150880696
Šikmost	-0,3053649307	0,3125474254
Odchylka od 0 :	Nevýznamná	Nevýznamná
Špičatost :	2,434718754	1,484697937
Odchylka od 3 :	Nevýznamná	Nevýznamná
Modus :	2,673090909	2,372272727

Klasické parametry pro náš výzkum jakosti povrchu dezénované formy nemají velký význam, na data musíme pohlížet jako na data párová. Zajímá nás tedy diference d_i .

Homogenita, normalita i heteroskedasticita dat v jednotlivých sloupcích (Dezén nové formy, Dezén formy po 500 000 cyklech) není prakticky zajímavá.

Tab. 1.2 Párový test:

Hladina významnosti :	0,05	
	Dezén nové formy Rz (um)	Dezén formy po 500 000 cyklech Rz (um)
Test významnosti rozdílu		
Závěr :	Rozdíly jsou VÝZNAMNÉ	
Pravděpodobnost :	3,732969338E-005	

1.1 Statistický závěr

Dle tabulky Tab.1.2 (a amerického testování) Párového testu vidíme, že:

Pravděpodobnost $p = 3,732969338E-005 < \alpha = 0,050$ a je tedy možné H_0 zamítnout.

1.2 Praktický závěr

Na základě Tab. 1.2 můžeme konstatovat, že difference Rz dezénu formy nové a po 50 000 cyklech se liší statisticky významně ($d_{Rz} \neq 0$). Je tedy třeba formu opětovně podrobit technologii dezénování.

2.1. Stejný problém řešený programem Minitab 14

Paired T-Test and CI: Dezen_nove_formy; Dezen_formy_po_500000_cyklech

Tab.2.1 Párový test dat

Paired T for Dezen_nove_formy - Dezen_formy_po_500000_cyklech

	N	Mean	StDev	SE Mean
Dezen_nove_formy	10	2,65200	0,06730	0,02128
Dezen_formy_po_5	10	2,42500	0,06151	0,01945
Difference	10	0,227000	0,104780	0,033134

95% CI for mean difference: (0,152045; 0,301955)

T-Test of mean difference = 0 (vs not = 0): T-Value = 6,85 P-Value = 0,000

1.2 Statistický závěr

Dle tabulky Tab.2.1 (a amerického testování) Párového testu vidíme, že:

Pravděpodobnost $p = 0,000 < \alpha = 0,050$ a je tedy možné H_0 zamítnout.

1.2 Praktický závěr

Na základě Tab. 2.1 můžeme opětovně konstatovat, že difference Rz dezénu formy nové a po 500 000 cyklech se liší statisticky významně ($d_{Rz} \neq 0$). Je tedy třeba formu opětovně podrobit technologii dezénování.

Jednofaktorová ANOVA bez transformace

Zadání:

Nástrojový materiál ocel 19 600 byl strojně leštěn čtyřmi technologiemi, označenými A,B,C a D (úrovně faktoru A). Každý materiál byl po skončení leštění proměřen z hlediska jakosti povrchu parametrem Ra. Celkově se každé měření opakovalo 30krát. Cílem je posoudit vliv technologie na výsledek analýzy.

Data

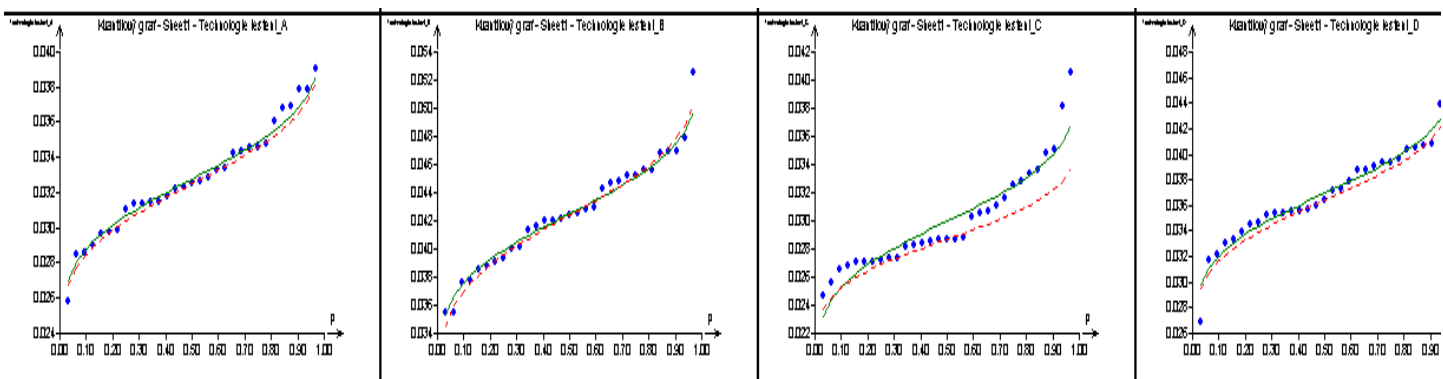
Data tj. parametry Ra byla měřena pro technologie označené A, B, C, a D (představující úrovně faktoru A).

Faktor A zahrnující vliv technologií leštění A, B, C a D. / Parametry Ra jsou v um										
TechnologieA	0,0346	0,0344	0,0299	0,0298	0,0379	0,0315	0,0334	0,0258	0,0325	0,0379
TechnologieA	0,0443	0,0386	0,0400	0,0414	0,0425	0,0390	0,0376	0,0428	0,0420	0,0429
TechnologieA	0,0381	0,0281	0,0348	0,0287	0,0268	0,0274	0,0272	0,0270	0,0287	0,0256
TechnologieB	0,0463	0,0372	0,0334	0,0379	0,0345	0,0387	0,0394	0,0356	0,0439	0,0354
TechnologieB	0,0346	0,0344	0,0299	0,0298	0,0379	0,0315	0,0334	0,0258	0,0325	0,0379
TechnologieB	0,0443	0,0386	0,0400	0,0414	0,0425	0,0390	0,0376	0,0428	0,0420	0,0429
TechnologieC	0,0381	0,0281	0,0348	0,0287	0,0268	0,0274	0,0272	0,0270	0,0287	0,0256
TechnologieC	0,0463	0,0372	0,0334	0,0379	0,0345	0,0387	0,0394	0,0356	0,0439	0,0354
TechnologieC	0,0346	0,0344	0,0299	0,0298	0,0379	0,0315	0,0334	0,0258	0,0325	0,0379
TechnologieD	0,0443	0,0386	0,0400	0,0414	0,0425	0,0390	0,0376	0,0428	0,0420	0,0429
TechnologieD	0,0381	0,0281	0,0348	0,0287	0,0268	0,0274	0,0272	0,0270	0,0287	0,0256
TechnologieD	0,0463	0,0372	0,0334	0,0379	0,0345	0,0387	0,0394	0,0356	0,0439	0,0354

Užitý program: QC-Expert 2.9 CZ, Adstat125-Student, MINITAB 14

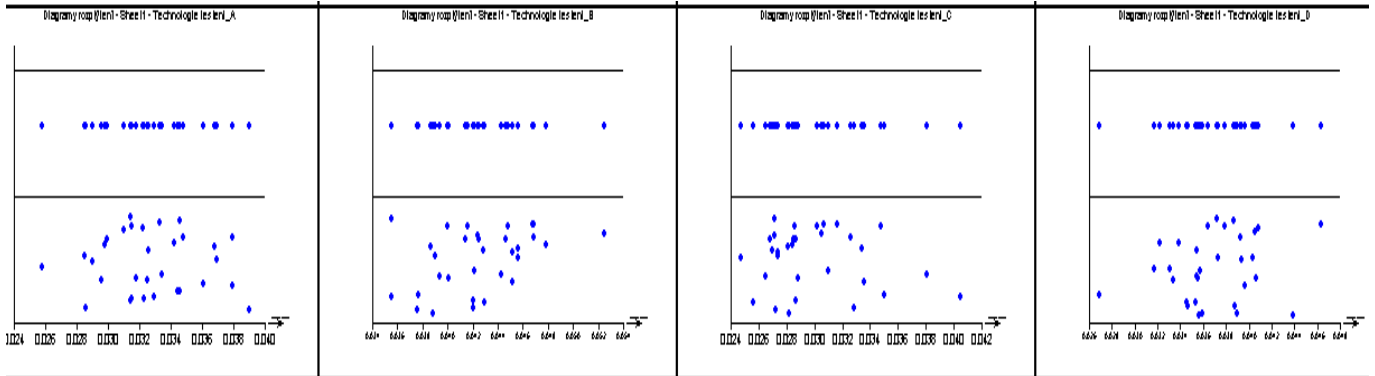
1. Ověření výběrových předpokladů

Obr.2.1 Kvantilové grafy



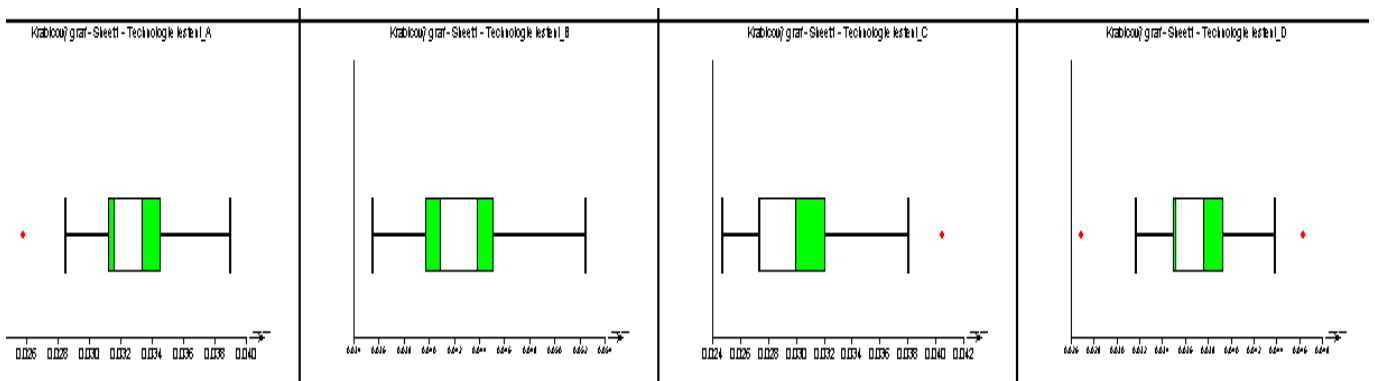
Kvantilové grafy pro technologie leštění A,B,C a D v případech A,B, a D ukazují na nepříliš výraznou datovou asymetrii, v případě faktorů A, B, C a D detekují přítomnost odlehlých hodnot.

Obr. 2.2 Bodové grafy



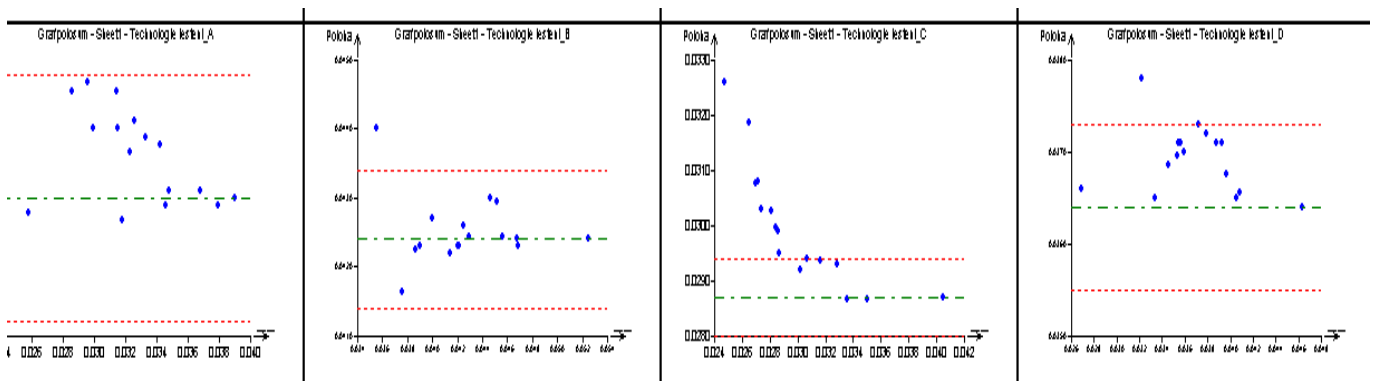
Diagramy rozptýlení a Rozmítnuté diagramy rozptýlení detekují mírnou asymetrii všech případů (technologí leštění), v případě technologií A, B, C a D přítomnost jednoho až dvou vybočených hodnot.

Obr. 2.3 Krabicové grafy



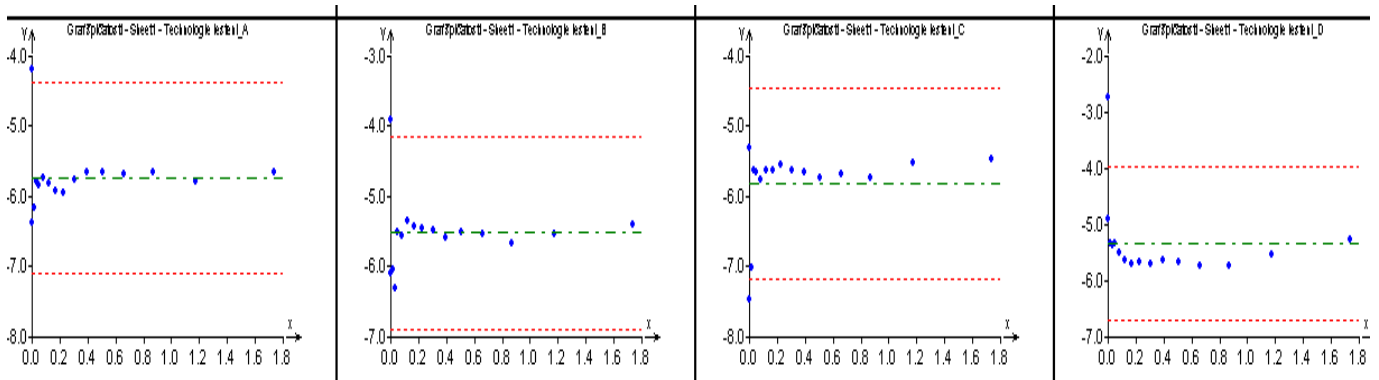
Taktéž krabicové grafy vykazují mírnou asymetrii a přítomnost vybočujících hodnot.

Obr. 2.4 Grafy polosum



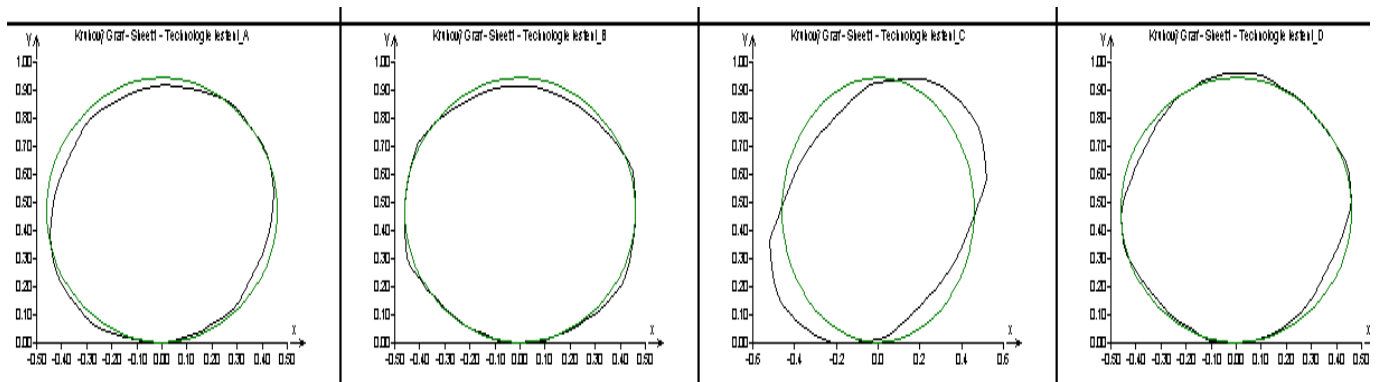
V grafech polosum mimo technologii C je patrná asymetrie rozdělení, především pro případ technologie C a částečně i pro případ D.

Obr. 2.5 Grafy špičatosti



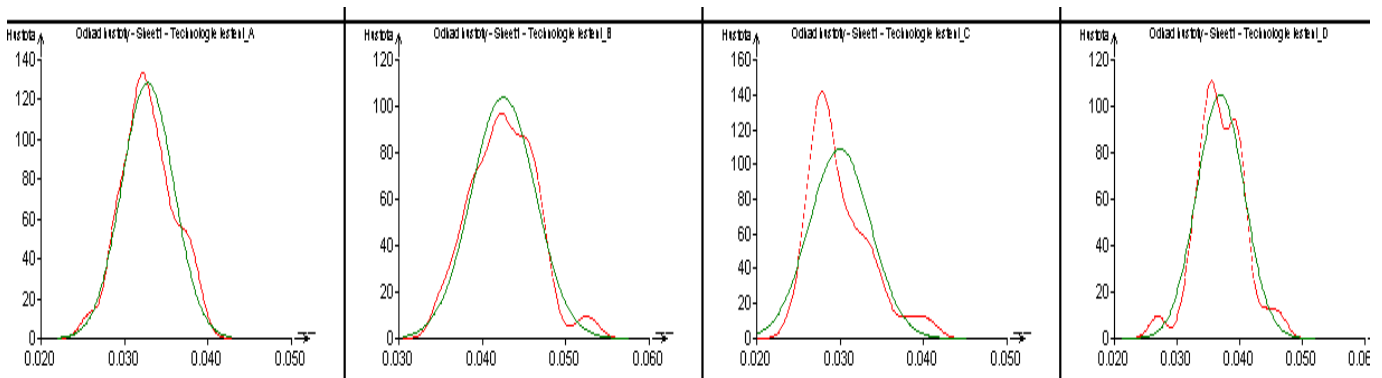
V grafech špičatosti je možné detekovat, že většina bodů leží rovnoběžně s osou x. Znamená to tedy, nevýraznou asymetrii naměřených dat.

Obr. 2.6 Kruhové grafy



V kruhových grafech je také patrná mírná asymetrie dat, poněkud problematicky se jeví především technologie C, u které je asymetrie větší, tedy směrem k vyšším hodnotám.

Obr. 2.7 Grafy odhadů hustot



Grafy jádrových odhadů hustot pravděpodobnosti detekují, že naměřená data jsou tvarově nepříliš odlišná od rozdělení Normálního. Odhady průměrů nejsou významně posunuty v důsledku vybočujících hodnot vůči odhadům mediánů.

Tab.2.1 Znaménkový test nezávislosti naměřených dat

Hladiny významnosti $\alpha = 0,05$

Znaménkový test				
Závěry	Data jsou nezávislá	Data jsou nezávislá	Data jsou nezávislá	Data jsou nezávislá

Dle znaménkového testu provedeném za účelem ověření předpokladů o výběrových souborech úrovně faktoru A (technologie leštění A, B, C a D), byla vypočtená nezávislost naměřených dat parametrů Ra.

Tab.2.1 Testy normality naměřených dat

Hladiny významnosti $\alpha = 0,05$

Test normality				
Název sloupce :	Technologie lesteni_A	Technologie lesteni_B	Technologie lesteni_C	Technologie lesteni_D
Průměr :	0,03278064516	0,04251935484	0,02998709677	0,03694193548
Rozptyl :	9,64227957E-006	1,47596129E-005	1,343049462E-005	1,441184946E-005
Šikmost	0,0717680817	0,2214806809	1,101284959	-0,01501518709
Špičatost :	2,624423251	2,97962265	3,810585911	3,789766538
Normalita :	Přijata	Přijata	Přijata	Přijata
Vypočtený :	0,1018863013	0,4912835915	5,789032409	0,05642481403
Teoretický :	5,991464547	5,991464547	5,991464547	5,991464547
Pravděpodobnost :	0,9503326948	0,782202363	0,05532578498	0,9721818466

Při provedení testů normalit naměřených dat lze konstatovat, že nezamítáme hypotézu normality dat (viz tab.2.1, řádek č.8 a 9) a Tab. 2.1.1. Nicméně technologie C a D má vyšší hodnotu špičatosti, technologie C i šikmosti.

Tab. 2.1.1 Porovnání vypočtené a tabulkové hodnoty t

	Technologie lesteni_A	Technologie lesteni_B	Technologie lesteni_C	Technologie lesteni_D
Vypočtená hodnota $t_{\text{vypočtená}}$	0,1018863013	0,4912835915	5,789032409	0,05642481403
Tabulková hodnota $t_{\text{tabulková}}$	5,991464547	5,991464547	5,991464547	5,991464547

Ve všech případech tabulky Tab.2.1.1 platí, že $t_{\text{vypočtená}} < t_{\text{tabulková}}$, a tedy nezamítáme hypotézu normality dat (v případě technologie C však jen těsně)

Tab.2.3 Testy vybočujících hodnot naměřených dat

Hladiny významnosti $\alpha = 0,05$

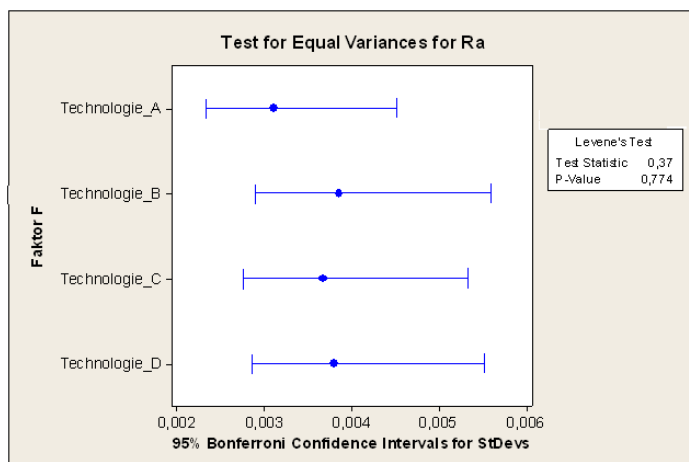
Vybočující body				
Název sloupce :	Technologie lesteni_A	Technologie lesteni_B	Technologie lesteni_C	Technologie lesteni_D
Homogenita :	Přijata	Přijata	Přijata	Přijata
Počet vybočujících bodů	0	0	0	0

Testy vybočujících hodnot naměřených dat na dané hladině významnosti nezamítají homogenitu naměřených dat úrovní faktoru A, můžeme tedy předpokládat homogenitu dat jednotlivých úrovní faktoru A.

Ověření Homoskedasticity naměřených dat úrovní faktoru A.

Hladiny významnosti $\alpha = 0,05$

Obr. 2.8 Ověření Homoskedasticity naměřených dat úrovní faktoru A



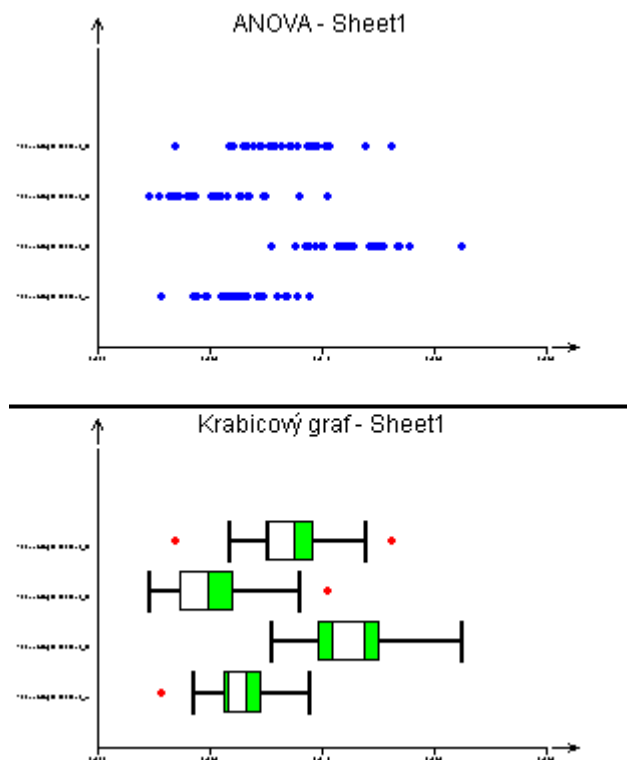
Jak plyne z Obr.2.8 , Levenuv test indikuje p hodnotu = 0,774 > $\alpha = 0.050$. Znamená to, že není možno zamítnout hypotézu o Homoskedasticitě naměřených dat úrovní faktoru A na zvolené hladině významnosti.

Závěr ověření výběrových předpokladů

Při hodnocení výběrových souborů bylo zjištěno, že data jsou nezávislá, pocházejí z normálního rozdělení, jsou homogenní a dodržují homoskedasticitu.

3. Aplikace jednofaktorové metody ANOVA bez transformace

Obr.3.1 Zobrazení poloh měřených dat



Graf ANOVA zobrazuje polohu měřených dat, takže je možné vizuálně posoudit rozdíly a rozptyl v jednotlivých úrovních faktoru F. Na ose x je měřená hodnota, tj. parametr Ra, na ose y jsou zadané úrovně faktoru A, tj. technologie leštění A,B,C a D. Krabicový graf je konstruován za stejným účelem.

Numerické výpočty a hypotézy:

Ho: Efekty faktoru A jsou nulové (tzn. že existuje nevýznamný vliv technologie leštění na Ra)

Ha: Efekty faktoru A nejsou nulové (tzn. že existuje významný vliv technologie leštění na Ra)

$$1-\alpha = 0,95$$

Tab.3.1 Parametry ANOVA získané softverem QCExpert

Celkový průměr :	0,03555725806		
Celkový rozptyl :	3,520474364E-005		
Průměrný čtverec :	3,492083442E-005		
Reziduální rozptyl :	1,274249672E-005		
Reziduální součet čtverců :	0,001567327097		
Celkový součet čtverců :	0,004330183468		
Vysvětlený součet čtverců :	0,002762856371		
Počet úrovní faktoru :	4		
Sloupec	Počet hodnot	Efekty faktorů	Průměr úrovně
Technologie lesteni_A	31	-2,7766129032258E-03	3,27806451612903E-02
Technologie lesteni_B	31	6,96209677419355E-03	4,25193548387097E-02
Technologie lesteni_C	31	-5,57016129032258E-03	2,99870967741935E-02
Technologie lesteni_D	31	1,38467741935483E-03	0,036941935483871
Test významnosti celkového vlivu faktoru :			
Závěr	Teoretický	Vypočítaný	Pravděpodobnost
Významný	2,68016757	70,51128961	2,332718688E-026

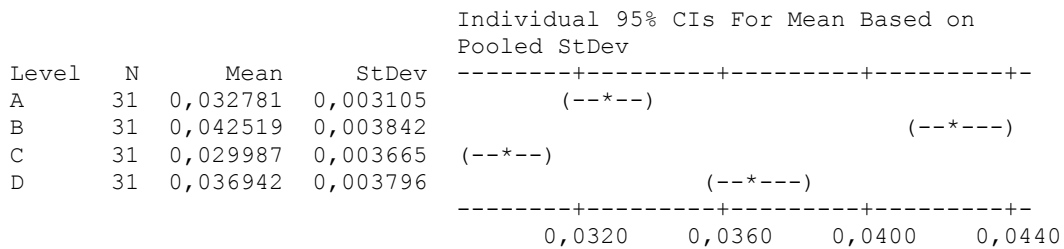
Z Tab. 3.1 plyne, že Fischer-Snedecorovo testační kritérium $F_{exp} = 70.51128961$ > než tabulkový kvantit $F_{1-\alpha} = 2.68016757$. Znamená to tedy, že je nutno zamítnou H_0 ve prospěch H_a (pochopitelně s uvažováním $\alpha = 0.050$ > $2,332718688E-026$). V praxi to tedy znamená, že vliv faktoru technologie leštění na hodnotu parametru R_a je statisticky významný.

Tab.3.2 Parametry ANOVA získané softverem Minitab 14

One-way ANOVA:

Zdroj	Stup.volnosti	SS	MS	F	P
Rozptylu		Součet čtverců	Prům.čtverec	Test Hl.výz	
Mezi úrovněmi	3	0,0027629	0,0009210	70,51	0,000
Rezidua	120	0,0015673	0,0000131		
Celkově	123	0,0043302			

S = 0,003614 R-Sq = 63,80% R-Sq(adj) = 62,90%



Pooled StDev = 0,003614

Stejných výsledků obdržíme i při aplikaci softveru Minitab 14. Pouze s rozdílem, že je zde aplikován americký způsob testování, tedy je přímo určena hodnota $p=0,000 < \alpha = 0,050$. Závěr je tedy stejný a to, že vliv technologie leštění na parametr Ra je statisticky významný.

Jednofaktorová ANOVA s transformací

Zadání

Tři frikční materiály používané na výrobu brzdících desek, označené A,B,C (faktor A) byly bezkontaktně proměřeny a pro každé měření byla určena hodnota Rq (um). Celkově se každé měření opakovalo 11krát, pro každý frikční materiál. Cílem bylo posouzení rozdílnosti frikčního materiálu (faktoru A) pomocí amplitudového parametru jakosti povrchu Rq.

1. Data

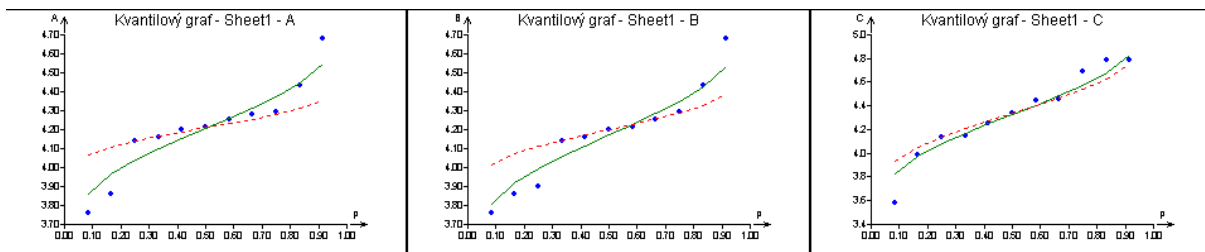
Data byla měřena pro materiály označené A, B, C (představující úrovně faktoru A) a velikosti výběru n = 11 pro každý z nich. Data jsou parametry Rq (um).

Faktor A zahrnující vliv frikčního materiálu A, B, C. / Parametry Rq jsou v um											
A	3,76	4,28	4,20	4,29	4,21	4,68	4,16	4,43	3,86	4,14	4,25
B	4,20	4,29	4,21	4,68	4,16	4,43	3,86	4,14	4,25	3,90	3,76
C	4,46	4,44	3,99	4,79	4,15	4,34	3,58	4,25	4,79	4,69	4,14

Užitý program: QC-Expert 2.9 CZ, Adstat125-Student, MINITAB 14

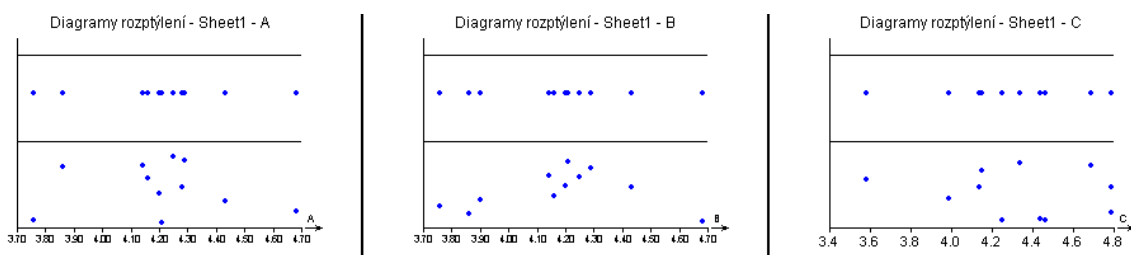
2. Ověření výběrových předpokladů

Obr.2.1 Kvantilové grafy



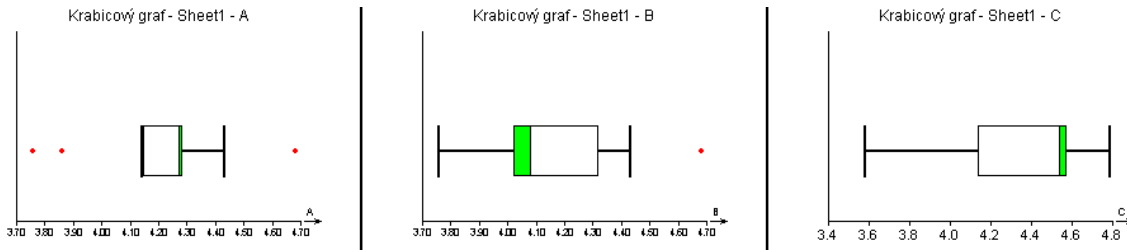
Kvantilové grafy pro frikční materiály A,B,C v případech A,B, ukazují na datovou asymetrii, v případě A, B, ale i C detekují přítomnost vybočujících hodnot.

Obr. 2.2 Bodové grafy



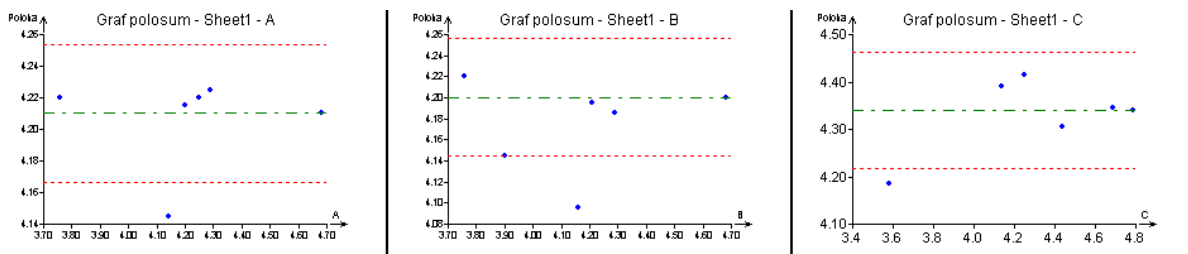
Diagramy rozptýlení a Rozmítnuté diagramy rozptýlení detekují datovou asymetrii a detekují i přítomnost vybočujících hodnot.

Obr. 2.3 Krabicové grafy



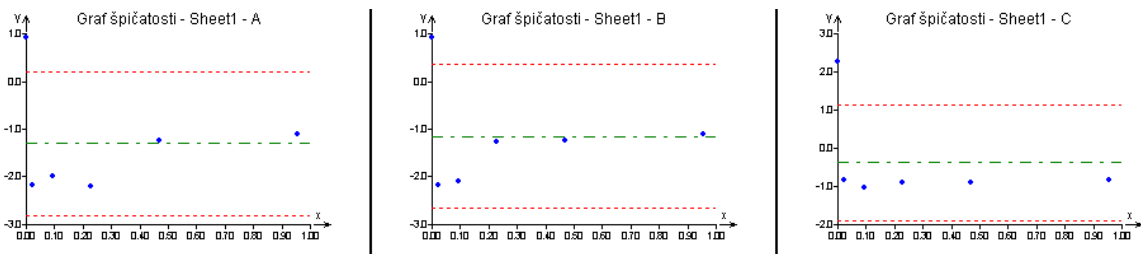
Taktěž krabicové grafy vykazují datovou asymetrii a přítomnost vybočujících hodnot.

Obr. 2.4 Grafy polosum



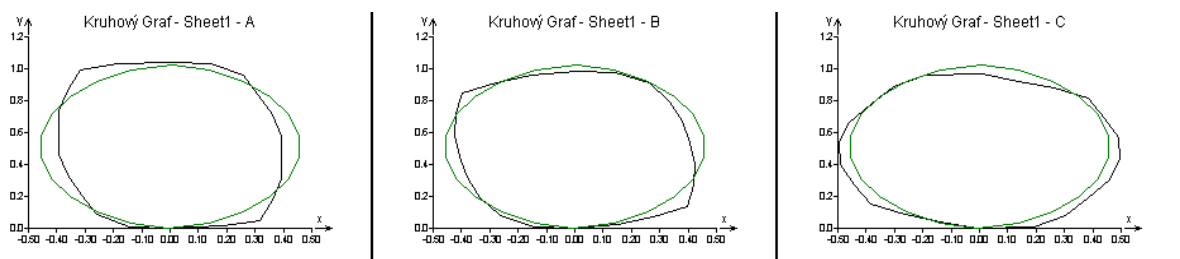
V grafech polosum je patrná datová asymetrie rozdělení. Dále je možné specifikovat, že zoom v ose y je relativně malý.

Obr. 2.5 Grafy špičatosti



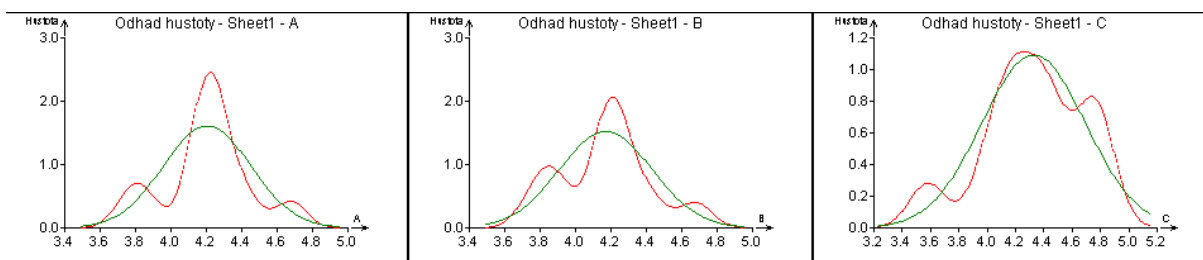
V grafech špičatosti je možné detekovat, že většina bodů leží rovnoběžně s osou x. Znamená to tedy, symetrii naměřených dat.

Obr. 2.6 Kruhové grafy



V kruhových grafech je patrná mírná asymetrie dat.

Obr. 2.7 Grafy odhadů hustot



Grafy jádrových odhadů hustot pravděpodobnosti detekují, že naměřená data jsou tvarově nepříliš odlišná od rozdělení Normálního. Toto tvrzení se však opírá o velmi malé soubory naměřených dat úrovní faktoru A. Dále v grafech můžeme vidět vliv vybočujících hodnot.

Tab.2.1 Znaménkový test nezávislosti naměřených dat

Velikosti výběrových souboru $n = 11$

Hladiny významnosti $\alpha = 0,05$

Znaménkový test	Frikční materiál A	Frikční materiál B	Frikční materiál C
Závěry	Data jsou nezávislá	Data jsou nezávislá	Data jsou nezávislá

Dle znaménkového testu provedeném za účelem ověření předpokladů o výběrových souborech úrovní faktoru A (typu frikčního materiálu A, B, C), byla vypočtená nezávislost naměřených dat parametrů R_q .

Tab.2.1 Testy normality naměřených dat

Velikosti výběrových souboru $n = 31$

Hladiny významnosti $\alpha = 0,05$

Test normality	Frikční materiál A	Frikční materiál B	Frikční materiál C
Název sloupce :	Frikční materiál A	Frikční materiál B	Frikční materiál C
Průměr :	4,205454545	4,170909091	4,329090909
Rozptyl :	0,06144727273	0,06890909091	0,1334890909
Šikmost	-0,07366771942	0,1870197184	-0,4857511977
Špičatost :	3,116634424	2,667745011	2,704500766
Normalita :	Přijata	Přijata	Přijata
Vypočtený :	0,06902112757	0,1965593059	0,987707028
Teoretický :	5,991464547	5,991464547	5,991464547
Pravděpodobnost :	0,9660781342	0,9063953922	0,6102701726

Při provedení testů normalit naměřených dat úrovní faktoru A lze konstatovat, že není možné zamítnout hypotézu, že data pocházejí ze souborů s normálním rozdělením (viz tab.2.1, řádek č.8 a 9) a Tab. 2.1.1.

Tab. 2.1.1 Porovnání vypočtené a tabulkové hodnoty t

	Frikční materiál _A	Frikční materiál _B	Frikční materiál _C
Vypočtená hodnota $t_{\text{vypočtená}}$	0,06902112757	0,1965593059	0,987707028
Tabulková hodnota $t_{\text{tabulková}}$	5,991464547	5,991464547	5,991464547

Ve všech případech tabulky Tab.2.1.1 platí, že $t_{\text{vypočtená}} < t_{\text{tabulková}}$, a tedy platí závěr o nemožnosti zamítnutí normálního rozdělení. Konstatujeme tedy, že data (parametr Rq) úrovní faktoru A (frikční materiál A, B, C) pochází ze souboru s normálním rozdělením.

Tab.2.3 Testy vybočujících hodnot naměřených dat

Velikosti výběrových souborů $n = 11$

Hladiny významnosti $\alpha = 0,05$

Vybočující body			
Název sloupce :	Frikční materiál _A	Frikční materiál _B	Frikční materiál _C
Homogenita :	Přijata	Přijata	Přijata
Počet vybočujících bodů	0	0	0

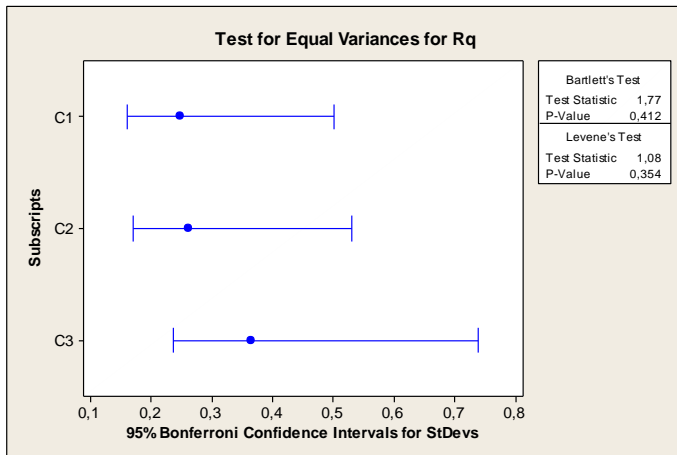
Testy vybočujících hodnot naměřených dat na dané hladině významnosti nezamítají homogenitu naměřených dat úrovní faktoru A, můžeme tedy předpokládat homogenitu naměřených dat úrovní faktoru A.

Ověření Homoskedasticity naměřených dat úrovní faktoru A.

Velikosti výběrových souborů $n = 11$

Hladina významnosti $\alpha = 0,05$

Obr. 2.8 Ověření Homoskedasticity naměřených dat úrovní faktoru A



Jak plyne z Obr.2.8 , Levenuv test indikuje p hodnotu = 0,354 $>$ $\alpha = 0.050$. Znamená to, že není možno zamítnout hypotézu o Homoskedasticitě naměřených dat úrovní faktoru A na dané hladině významnosti.

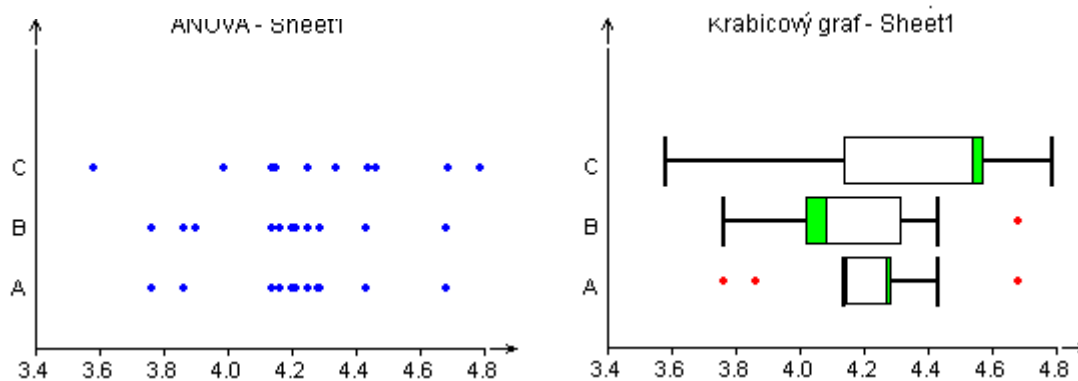
Stejného výsledku dosáhneme i při využití Bartlettova testu, kde p hodnota = 0,412 $>$ $\alpha = 0.050$.

Závěr ověření výběrových předpokladů

Při hodnocení výběrových souborů o $n = 11$ parametru Rq bylo zjištěno, že data jsou nezávislá, pocházejí z normálního rozdělení, jsou homogenní a dodržují homoskedasticitu. Nicméně je nutno uvážit, že velikost výběrových souborů $n=11$, což ovlivňuje sílu jednotlivých metod.

3. Aplikace jednofaktorové metody ANOVA bez a s transformací

Obr.3.1 Zobrazení poloh měřených dat



Graf ANOVA zobrazuje polohu měřených dat, takže je možné vizuálně posoudit rozdíly a rozptyl v jednotlivých úrovních faktoru A. Na ose x je měřená hodnota, tj. parametr Rq, na ose y jsou zadané úrovně faktoru A, tj. frikční materiál A,B a C. Krabicový graf je konstruován za stejným účelem.

Numerické výpočty a hypotézy:

Ho: Efekty faktoru A jsou nulové (tzn. že neexistuje významný vliv frikčního materiálu na Rq)

Ha: Efekty faktoru A nejsou nulové (tzn. že existuje významný vliv frikčního materiálu na Rq)

$$1-\alpha = 0,95$$

Tab.3.1 Parametry ANOVA získané softverem ADStat student

Jednofaktorová ANOVA bez logaritmické transformace

TABULKA ANOVA:

H0: Efekty faktoru A jsou nulové, HA: ... nejsou nulové

Kvantil $F(1-\alpha, k-1, n-k) = 3.316$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H0 je	Spočtená hlad.výz.
Mezi úrovněmi	k-1 = 2	1.5217E-01	7.6085E-02	0.865	Akceptována	0.431
Rezidua	n-k = 30	2.6385E+00	8.7948E-02			
Celkový	n-1 = 32	2.7906E+00	8.7207E-02			

Jednofaktorová ANOVA s logaritmickou transformací

TABULKA ANOVA:

H0: Efekty faktoru A jsou nulové, HA: ... nejsou nulové

Kvantil $F(1-\alpha, k-1, n-k) = 3.316$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H0 je	Spočtená hlad.výz.
Mezi úrovněmi	k-1 = 2	7.6416E-03	3.8208E-03	0.765	Akceptována	0.474
Rezidua	n-k = 30	1.4981E-01	4.9937E-03			
Celkový	n-1 = 32	1.5745E-01	4.9204E-03			

Závěr statistický

Z Tab. 3.1 plyne, že dle amerického testování je spočtená hladina významnosti $p = 0,431 > \alpha = 0,050$. Obdobného výsledku dosáhneme v případě transformace, tj. $p = 0,474 > \alpha = 0,050$. Znamená to tedy, že není možné zamítnout Ho, tj. že efekty faktoru A jsou nulové.

Závěr praktický

Cílem bylo posouzení rozdílnosti frikčních materiálů (faktoru A). Můžeme konstatovat, že s $1-\alpha = 0,95$ není rozdíl mezi frikčními materiály A,B a C významný, z hlediska naměřeného parametru Rq.

Dvoufaktorova ANOVA bez opakování:

Zadání:

Zadání: Vzorok materiálu formy pro vstřikování plastů byly leštěny čtyřmi techniky (kteří používali stejný předepsaný technologický postup). Každý vyleštěný vzorek byl potom přeměřen z hlediska jakosti povrchu a stanoven parametr Rt pěti profilometry, které se lišily dle výrobce. Pomocí analýzy rozptylu je nutné zjistit, zda na jakost leštěného povrchu má statisticky významný vliv typ profilometru a technik provádějící leštění.

Data:

Faktor B	Tech.leštění_1	Tech.leštění_2	Tech.leštění_3	Tech.leštění_4
Faktor A	Rt (um)	Rt (um)	Rt (um)	Rt (um)
1_Přístroj	2,09	2,04	1,99	2,19
2_Přístroj	2,50	2,62	2,70	2,48
3_Přístroj	2,55	2,31	2,15	2,44
4_Přístroj	2,48	2,12	2,35	2,57
5_Přístroj	1,96	2,12	2,21	2,11

Užitý program: Adstat125-Student

1. ANOVA bez transformace

1.1 Podmínky:

Hladina významnosti $\alpha = 0,05$

Transformace Ne

Metoda analýzy Nejmenší čtverce

1.2 Stanovení hypotéz průměrů a efektů:

Faktor A přesnost přístroje dle výrobce

Faktor B technik provádějící leštění

Ho: Efekty faktoru A jsou nulové (přesnost přístroje dle výrobce nemá vliv na měření Rt)

Ha: Efekty faktoru A nejsou nulové (přesnost přístroje dle výrobce má vliv na měření Rt)

Ho: Efekty faktoru B jsou nulové (technik provádějící leštění nemá vliv na jakost povrchu)

Ha: Efekty faktoru B nejsou nulové (technik provádějící leštění má vliv na jakost povrchu)

Ho: Interakce faktorů A a B neexistuje

Ha: Interakce faktorů A a B existuje

PRŮMĚRY A EFEKTY ÚROVNÍ:

Celkový průměr = 2.2990E+00
Reziduální rozptyl = 2.3014E-02

F A K T O R A:			F A K T O R B:		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	2.0775E+00	-2.2150E-01	1	2.3160E+00	1.7000E-02
2	2.5750E+00	2.7600E-01	2	2.2420E+00	-5.7000E-02
3	2.3625E+00	6.3500E-02	3	2.2800E+00	-1.9000E-02
4	2.3800E+00	8.1000E-02	4	2.3580E+00	5.9000E-02
5	2.1000E+00	-1.9900E-01			

Tukeyho C = -1.0743E+00

1.3 Anova pro model s Tukeyho iteracemi:

TABULKA ANOVA PRO MODEL S TUKEYHO INTERAKCÍ:

H0: Efekty faktoru A jsou nulové, HA: ... nejsou nulové

Kvantil $F(1-\alpha, n-1, mn-m-n) = 3.357$

H0: Efekty faktoru B jsou nulové, HA: ... nejsou nulové

Kvantil $F(1-\alpha, m-1, mn-m-n) = 3.587$

H0: Interakce I je nulová, HA: ... není nulová

Kvantil $F(1-\alpha, 1, mn-m-n) = 4.844$

(Zde I znamená efekt Tukeyho interakce.)

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H0 je	Spočtená hlad. výz.
Mezi úrovněmi A	n-1 = 4	7.0173E-01	1.7543E-01	7.623	Zamítnuta	0.003
Mezi levels B	m-1 = 3	3.6900E-02	1.2300E-02	0.534	Akceptována	0.668
Interakce	1	1.4941E-03	1.4941E-03	0.065	Akceptována	0.804
Rezidua	mn-m-n = 11	2.5315E-01	2.3014E-02			
Celkový	mn-1 = 19	9.9178E-01	5.2199E-02			

1.3.1 Statistický závěr:

Protože vypočtená hladina významnosti $p = 0,003 < \alpha = 0,050$ je nulová hypotéza H_0 zamítnuta a faktor A přesnost přístroje (profilometru) daná výrobcem je statisticky významná.

Protože vypočtená hladina významnosti $p = 0,668 > \alpha = 0,050$ je nulová hypotéza H_0 přijata a faktor B Vliv leštícího technika je statisticky nevýznamný.

Protože vypočtená hladina významnosti $p = 0,804 > \alpha = 0,050$ je nulová hypotéza H_0 přijata a interakce faktorů A Přesnosti měřícího přístroje a B Leštící technik je statisticky nevýznamný.

1.3.2 Technický závěr:

Vliv měřícího přístroje (profilometru) z hlediska přesnosti dané výrobcem má významný vliv na posouzení jakosti povrchu formy z hlediska parametru R_t .

Vlivy technika a interakce přesnost měřícího přístroje a technika nebyly prokázány

1.4 Zkouška transformace:

Odhad mocninné transformace : 3.4697E+00
Rozptyl odhadu transformace : 8.6121E+01
Akceptovatelný interval : (-5.8104E+00, 1.2750E+01)

Odhad mocninné transformace (3,4697) leží v akceptovatelném intervalu (-5,8104 12,75), proto není třeba na data aplikovat mocninnou transformaci.

Dvufaktorová ANOVA s opakování

Zadání:

Při tvorbě desénu pro formy na vstřikování polymerního materiálu ABS byl vzorek mající stejné vlastnosti jako forma pískována po dobu 20, 60, 90 a 120 min. K tomu byly použity pískovací směsi o složení 1,2,3 a 4. Po provedení pískování formy předepsanou dobou a směsí byl hodnocen kontaktním způsobem (profiloměrem Taylor Hobson INTRA) parametrem jakosti povrchu Rz (který byl automaticky softverem profiloměru zaokrouhlen na celá čísla). Toto se opakovalo pro konkrétní dobu a směs 5x.

Úkolem je zjistit, zda doba pískování a složení pískovací směsi je významná pro požadovanou hodnotu jakosti parametru Rz povrchu.

Data:

Doba pískování (min)	Složení pískovací směsi_1	Složení pískovací směsi_2	Složení pískovací směsi_3
	Faktor B	Faktor B	Faktor B
20 min. Faktor A	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)
	5	1	3
	6	2	1
	9	1	3
	5	2	3
6	1	4	
60 min. Faktor A	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)
	11	1	4
	10	3	4
	9	2	5
	9	3	4
10	2	5	
90 min. Faktor A	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)
	12	3	6
	10	4	4
	13	2	5
	14	4	5
14	3	6	
120 min Faktor A	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)	Parametr Rz (um) (zaokrouh.na celá čísla)
	13	5	4
	12	4	5
	12	6	6
	13	3	6
13	4	4	

Užitý program: Adstat125-Student

1. ANOVA bez transformace

1.1 Podmínky:

Hladina významnosti	$\alpha = 0,05$
Transformace	Ne
Počet úrovní faktoru A „Doba pískování“	4
Počet úrovní faktoru B „Složení pískovací směsi“	3
Počet opakování v jedné buňce	5

1.3 Stanovení hypotéz průměrů a efektů:

Faktor A Doba pískování

Ho: Efekty faktoru A jsou nulové (efekt doby pískování na jakost povrchu je nevýznamný)

Ha: Efekty faktoru A nejsou nulové (efekt doby pískování na jakost povrchu je významný)

Faktor B Složení pískovací směsi

Ho: Efekty faktoru B jsou nulové (efekt složení pískovací směsi na jakost povrchu je nevýznamný)

Ha: Efekty faktoru B nejsou nulové (efekt složení pískovací směsi na jakost povrchu je významný)

Interakce faktoru A Doba pískování a B Složení pískovací směsi

Ho: Interakce faktorů A a B neexistuje

Ha: Interakce faktorů A a B existuje

Celkový průměr = 5.8167E+00
Reziduální rozptyl = 1.0583E+00

F A K T O R A:			F A K T O R B:		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	3.4667E+00	-2.3500E+00	1	1.0300E+01	4.4833E+00
2	5.4667E+00	-3.5000E-01	2	2.8000E+00	-3.0167E+00
3	7.0000E+00	1.1833E+00	3	4.3500E+00	-1.4667E+00
4	7.3333E+00	1.5167E+00			

1.3 Anova pro model s iteracemi faktoru A Doba pískování a faktoru B Složení pískovací směsi:

V tomto případě byl použit model s iteracemi faktorů **Doba pískování** a **Složení pískovací směsi**, protože je možné, že mezi vyjmenovanými faktory by mohl existovat vztah.

TABULKA ANOVA PRO MODEL S INTERAKCEMI FAKTORŮ A, B:

H₀: Efekty faktoru A jsou nulové, H_A: ... nejsou nulové

Kvantil $F(1-\alpha, n-1, mn(o-1)) = 2.798$

H₀: Efekty faktoru B jsou nulové, H_A: ... nejsou nulové

Kvantil $F(1-\alpha, m-1, mn(o-1)) = 3.191$

H₀: Interakce I je nulová, H_A: ... není nulová

Kvantil $F(1-\alpha, (n-1)(m-1), mn(o-1)) = 2.295$

(Zde I znamená efekty interakcí A a B dohromady)

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H ₀ je	Spočtená hlad.výz.
Mezi úrovněmi A	n-1 = 3	1.4018E+02	4.6728E+01	44.152	Zamítnuta	0.000
Mezi úrovněmi B	m-1 = 2	6.2703E+02	3.1352E+02	296.236	Zamítnuta	0.000
Interakce (n-1)(m-1)	= 6	4.0967E+01	6.8278E+00	6.451	Zamítnuta	0.000
Rezidua	mn(o-1) = 36	5.0800E+01	1.0583E+00			
Celkový	mno-1 = 59	8.5898E+02	1.4559E+01			

1.3.1 Statistický závěr:

Vyšetření vlivu faktoru A Doba pískování

Protože dle amerického testování je p hodnota = $0 < \alpha = 0,050$, je nulová hypotéza H₀ zamítnuta a faktor **Doba pískování** na parametr jakosti povrchu Rz je statisticky významný.

Vyšetření vlivu B faktoru Složení pískovací směsi

Protože dle amerického testování je p hodnota = $0 < \alpha = 0,050$, je nulová hypotéza H₀ zamítnuta a faktor **Složení pískovací směsi** na parametr jakosti povrchu Rz je statisticky významný.

Vyšetření vlivu interakce Doba pískování a Složení pískovací směsi

Protože dle amerického testování je p hodnota = $0 < \alpha = 0,050$, je nulová hypotéza H₀ zamítnuta a interakce faktorů **Doba pískování** a **Složení pískovací směsi** na parametr jakosti povrchu Rz je statisticky významná.

1.4 Zkouška transformace

Korelační koeficient, R : 0.272

(Pokud je R blízké nule není transformace nutná.)

Korelační koeficient nabývá hodnoty 0,272, transformace bude aplikována.

1.5 Anova s logaritmicou transformací pro model s iteracemi faktorů Složení pískovací směsi a Doba pískování na parametr jakosti povrchu Rz:

Hladina významnosti $\alpha = 0,05$

Transformace Logaritmicá

Počet úrovní faktoru A „Doba pískování“	4
Počet úrovní faktoru B „Složení pískovací směsi“	3
Počet opakování v jedné buňce	5

TABULKA ANOVA PRO MODEL S INTERAKCEMI FAKTORŮ A, B:
H₀: Efekty faktoru A jsou nulové, H_A: ... nejsou nulové
Kvantil $F(1-\alpha, n-1, mn(o-1)) = 2.798$
H₀: Efekty faktoru B jsou nulové, H_A: ... nejsou nulové
Kvantil $F(1-\alpha, m-1, mn(o-1)) = 3.191$
H₀: Interakce I je nulová, H_A: ... není nulová
Kvantil $F(1-\alpha, (n-1)(m-1), nm(o-1)) = 2.295$
(Zde I znamená efekty interakcí A a B dohromady)

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H ₀ je	Spočtená hlad.výz.
Mezi úrovněmi A	n-1 = 3	6.6320E+00	2.2107E+00	27.768	Zamítnuta	0.000
Mezi úrovněmi B	m-1 = 2	1.9732E+01	9.8658E+00	123.923	Zamítnuta	0.000
Interakce (n-1)(m-1) = 6		6.2681E-01	1.0447E-01	1.312	Akceptována	0.270
Rezidua mn(o-1) = 36		3.8214E+00	7.9612E-02			
Celkový mno-1 = 59		3.0812E+01	5.2223E-01			

1.5.1 Statistický závěr:

Vyšetření vlivu faktoru A Doba pískování

Protože dle amerického testování je p hodnota = $0 < \alpha = 0,050$, je nulová hypotéza H₀ zamítnuta a faktor **Doba pískování** na parametr jakosti povrchu Rz je statisticky významný.

Vyšetření vlivu faktoru B Složení pískovací směsi

Protože dle amerického testování je p hodnota = $0 < \alpha = 0,050$, je nulová hypotéza H₀ zamítnuta a faktor **Složení pískovací směsi** na parametr jakosti povrchu Rz je statisticky významný.

Vyšetření vlivu interakce Doba pískování a Složení pískovací směsi

Protože dle amerického testování je p hodnota = $0,270 > \alpha = 0,050$, je nulová hypotéza H₀ nezamítnuta (akceptována) a interakce faktorů Složení pískovací směsi a Doba pískování na parametr jakosti povrchu Rz není statisticky významná.

1.6 Zkouška transformace

Korelační koeficient, R : -0.828
(Pokud je R blízké nule není transformace nutná.)

Korelační koeficient po logaritmické transformaci nabývá hodnoty -0,828 a před transformací 0,272. Můžeme tedy konstatovat, že rozdíl korelačního koeficientu před a po transformaci je významně rozdílný, což nás vede k tvrzení, že transformace byla nutná.

1.7 Závěr

Pomocí analýzy rozptylu bylo zjištěno, že na požadovanou jakost povrchu formy vyjádřenou parametrem Rz má vliv jak doba pískování, tak složení pískovací směsi.

Interakce faktorů Doba pískování a Složení pískovací směsi na parametr jakosti povrchu Rz nebyla prokázána.

Použitá literatura

1. MELOUN, Milan a Jiří MILITKÝ. *Statistická analýza experimentálních dat*. Vyd. 2. uprav. rozš. Praha: ACADEMIA, 2004, 953 s. ISBN 80-200-1254-0.
2. MELOUN, Milan. *Kompendium statistického zpracování dat: metody a řešené úlohy*. Vyd. 2., přeprac. a rozš. Praha: Academia, 2006, 982 s. ISBN 80-200-1396-2.
3. MELOUN, Milan. *Statistické zpracování experimentálních dat*. 2. vyd. Praha: Ars magna, 1998, 839 s. ISBN 80-721-9003-2