

Analýza kovariance

Aplikovaná statistika II.

Datový soubor

- Howellsova databáze
 - Databáze kranio-metrických údajů
 - uvedené v milimetrech (mm)
 - Počet populací: 30
 - Počet landmarku: 82
- Analyzovaný soubor
 - Population : Bushman
 - Landmarks: Výška obličeje - BPL
Bizygomatická šířka – ZYB

Analyzovaný soubor

	Rozsah	Výška obličeje		Bizygomatická šířka	
		Průměr	Směrodatná odchylka	Průměr	Směrodatná odchylka
Muži	48	93.659	5.285	123.561	4.691
Ženy	41	90.625	4.470	116.458	5.327
Celkem	89	92.022	5.068	119.730	6.151

Testované hypotézy

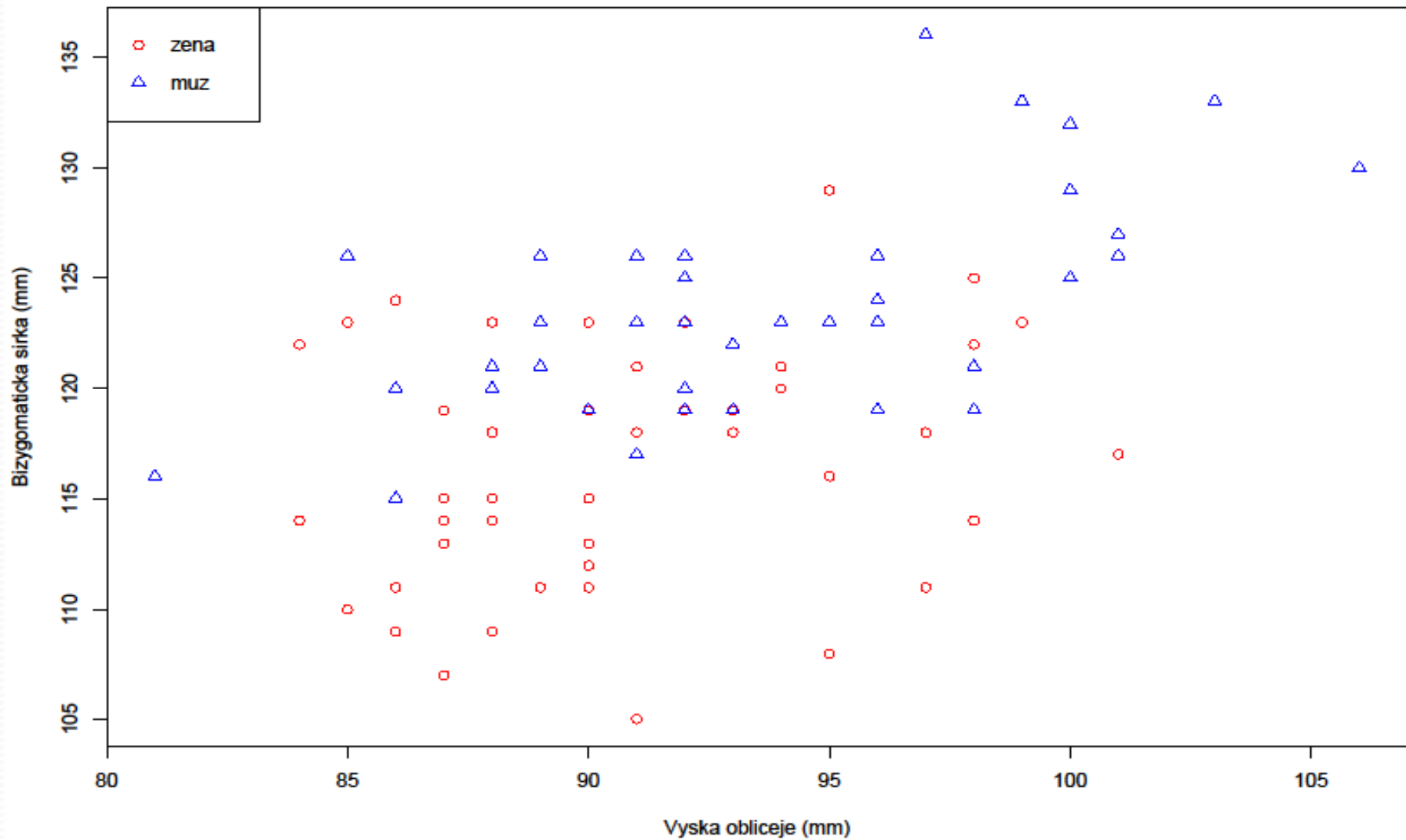
1. Hypotéza

- Ženy a muži populace Bushman se liší v bizygomatické šířce pokud kontrolujeme výšku obličeje

2. Hypotéza

- Faktor pohlaví je významný
- Použijeme analýzu kovariance
- Hypotézy testujeme na hladině významnosti 0.05

Rozložení dat



Ověření předpokladů

1. Normalita dat

- Shapirův-Wilkův test
 - $W = 0.98$; p-hodnota = 0.73

2. Nulová střední hodnota reziduí

- T-test
 - Statistika = blízka 0 ; p-hodnota = 1

3. Nezávislost reziduí

- Durbin-Watsonův test
 - Statistika = 1.87; p-hodnota = 0.44

4. Rovnost rozptylů

- Z grafického posouzení nevidíme problém

Ověření předpokladů

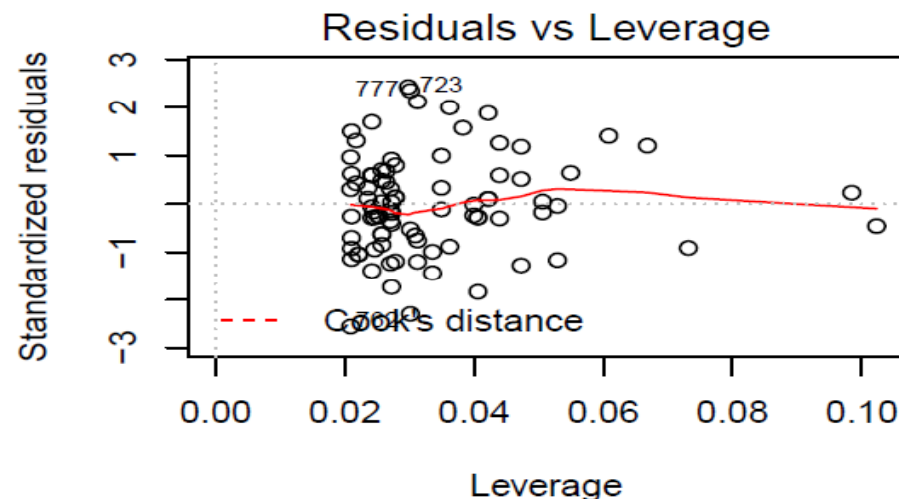
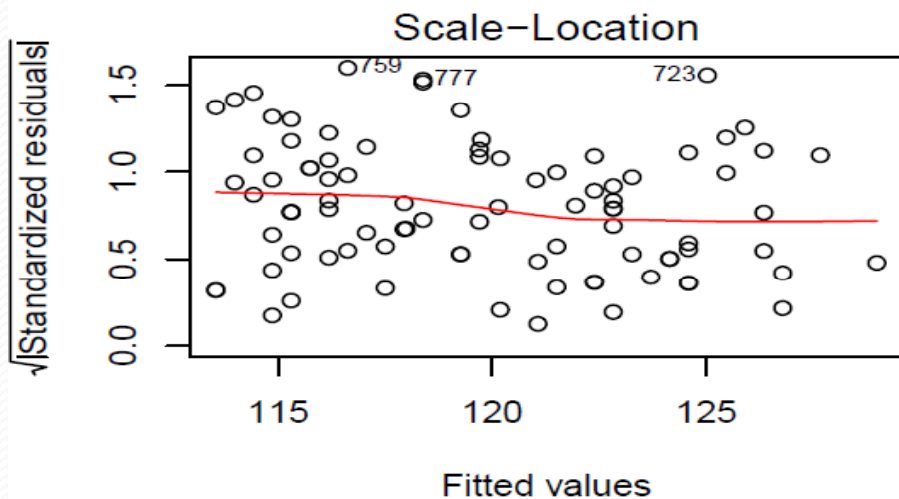
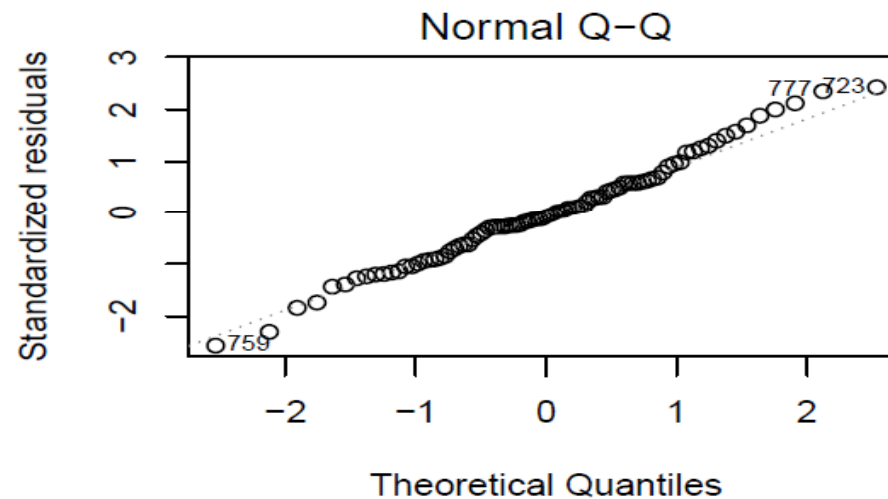
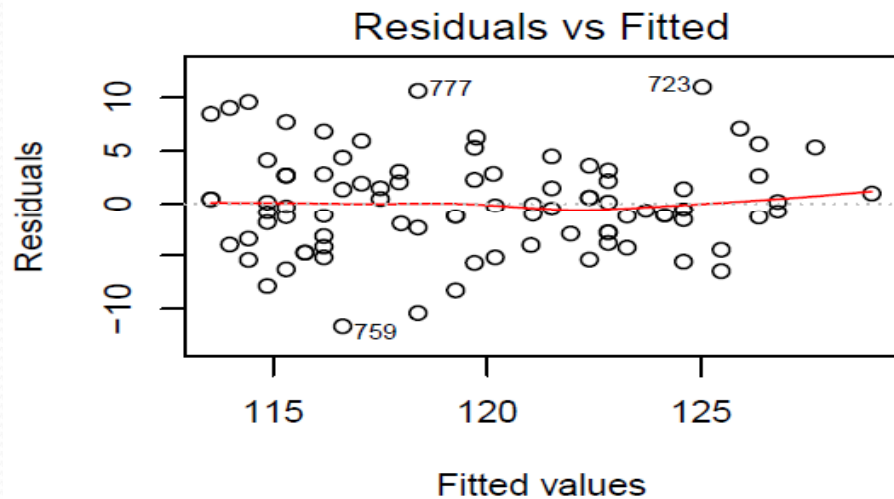
5. Rovnoběžnost regresních přímk

- Ověřujeme předpoklad, že regresní přímka modelující závislost výšky obličeje na bizygomatické šířce pro ženy je rovnoběžná s regresní přímkou pro muže
 - sestavíme model s interakcí (tj. model různoběžných přímk) a bez interakce (tj. model rovnoběžných přímk)
 - testujeme, zda je model bez interakce dostatečný
 - pro test využijeme funkci anova
-
- Hodnota testovací statistiky $F_{obs} = 1.46$
 - p -hodnota = 0.23
 - *Model bez interakce je dostatečný*
 - *Předpoklad rovnoběžnosti regresních přímk je splněný*

Ověření předpokladů

- Předpoklady modelu považujeme za splněny
- Shapirův-Wilkův test nabývá hodnoty 0.98 s p-hodnotou 0.73. Předpoklad normality považujeme za splněný
- T-test nabývá hodnoty blízké 0 s p – hodnotou blízkou 1. Hypotézu o nulové hodnotě reziduí nezamítáme
- Durbin-Watsonův test nabývá hodnoty 1.87 s p-hodnotou 0.44. Hypotézu o nezávislosti reziduí nezamítáme
- Předpoklad rovnosti rozptylů reziduí je na základě grafického posouzení v pořádku

Grafické posouzení



Významnost faktoru pohlaví

- Provedeme test shodnosti regresních přímek
 - Analýza rozptylu modelu bez faktoru pohlaví a modelu bez interakce
- Testujeme hypotézu, že regresní přímky jsou shodné
 - Hodnota testovací statistiky $F_{obs} = 31.70$
 - p -hodnota = blízka 0
 - Model s faktorem pohlaví je dostatečný
 - Faktor pohlaví je významný

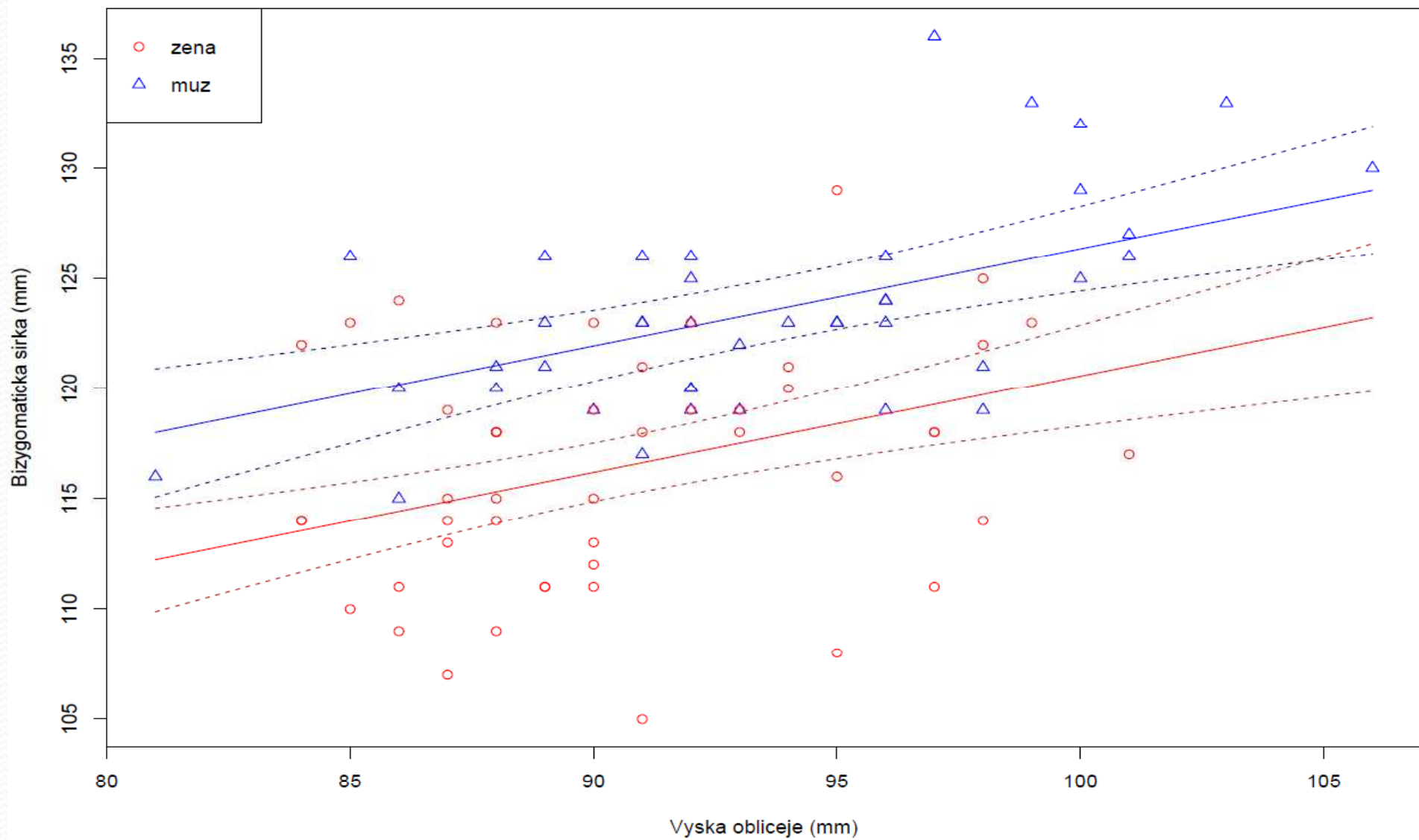
Odhad koeficientů a interpretace pro model bez interakce

- β_0 - odhad směrnice přímky pro kategorii ženy
 - Šířka obličeje u žen nabývá hodnoty 76.5720 pokud výška obličeje nabývá hodnoty 0
- β_1 - posun směrnice přímky pro kategorii muži
 - Šířka obličeje vzroste o 5.7675 mm pokud jedinec spadá do kategorie muži
- β_2 - společná směrnice přímky
 - Jestliže vzroste výška obličeje o 1mm, tak se zvětší šířka obličeje o 0.4401 mm
- Index determinace = 0.442
 - Variabilita v bizygomatické šířce mezi ženami a muži je z 44% vysvětlena výškou obličeje

Dílčí t-testy a F-test

- Dílčí t-testy
 - β_0 - t-test = 8.317; p – hodnota = blízka 0
 - β_1 - t-test = 5.631; p – hodnota = blízka 0
 - β_2 - t-test = 4.344; p – hodnota = blízka 0
 - Zamítáme nulové hypotézy o nevýznamnosti konstant na hladině významnosti 0.05
- F-test
 - Hodnota statistiky = 35.85; p – hodnota = blízka 0
 - Hypotézu o nevýznamnosti modelu jako celku zamítáme
 - Bizygomatickou šířku lze modelovat na výšce obličeje

Výsledný graf



Závěr

- Hypotéza, že muži a ženy populace Bushman se neliší v bazygomatické šířce byla zamítnuta
- Hypotéza, že faktor pohlaví je nevýznamný byla zamítnuta
- Variabilitu v bazygomatické šířce mezi ženami a muži populace Bushman vysvětluje z 44 % výška obličeje (náš model je přesný z 44%).



Děkuji za pozornost