



METODOLOGIJA NAUČNOG ISTRAŽIVANJA

DETEKCIJA I TRETMAN AUTLAJERA

Predrag Teovanović



UNIVERZITET U BEOGRADU
FAKULTET ZA SPECIJALNU EDUKACIJU I REHABILITACIJU
MASTER AKADEMSKE STUDIJE

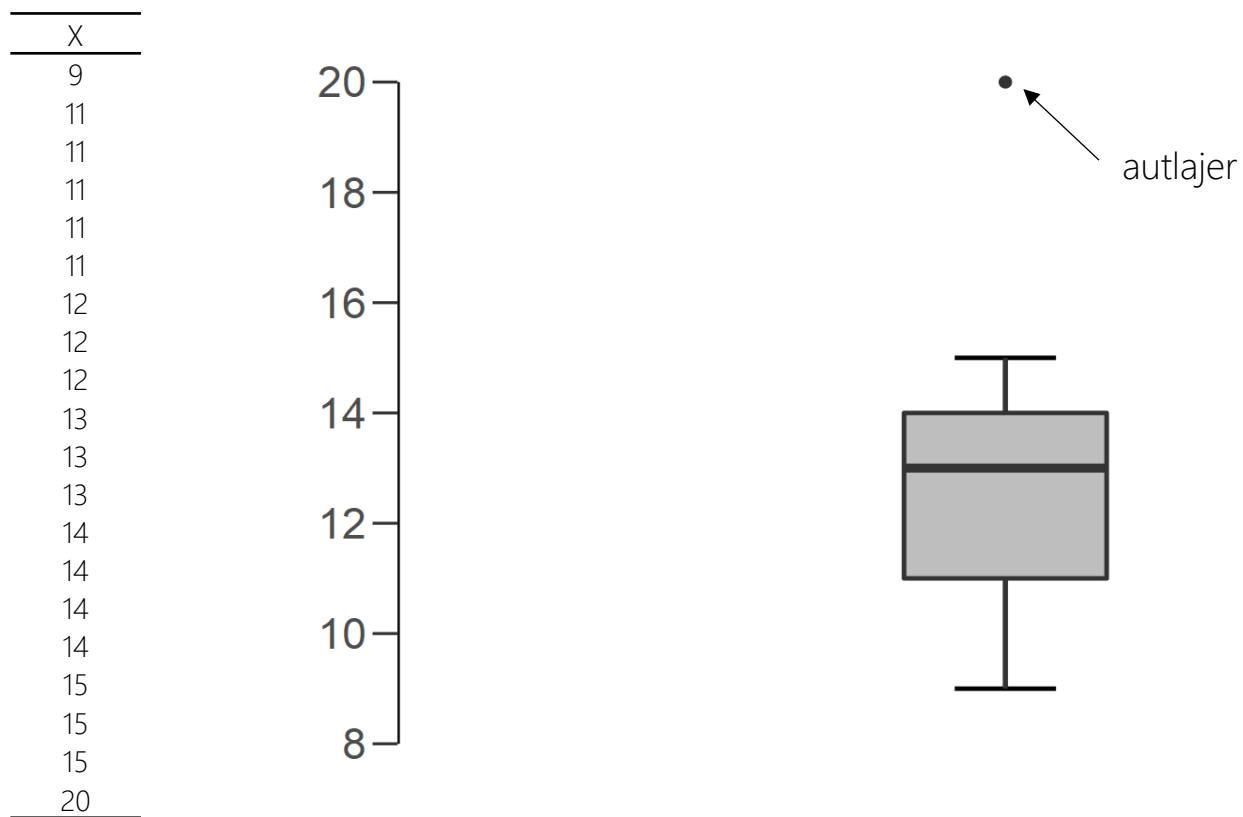


AUTLAJERI

Podaci koji se nalaze van raspona prihvatljivih vrednosti

Sinonimi: netipična tačka, nesaglasna vrednost, štrčak, iznimak, ekstremna vrednost

Autlajeri su ili slučajevi koji ne pripadaju datoj populaciji ili su posledica greške u unosu





TEŠKOĆE SA AUTLAJERIMA

Neke deskriptivne mere su naročito osetljive na pojavu autlajera (prosek, raspon, varijansa)

X	Statistik	Sa autlajerom	Bez autlajera
9	Prosek	13	12.6
11	Raspon	11 (20 – 9)	6 (20 – 15)
11	Varijansa	5.47	2.91
11	Standardna devijacija	2.43	1.70
12			
12			
12			
13			
13			
13			
14			
14			
14			
14			
15			
15			
15			
20			



TEŠKOĆE SA AUTLAJERIMA

Neke druge mere su manje osetljive na pojavu autlajera (mod, medijana, interkvartilni raspon)

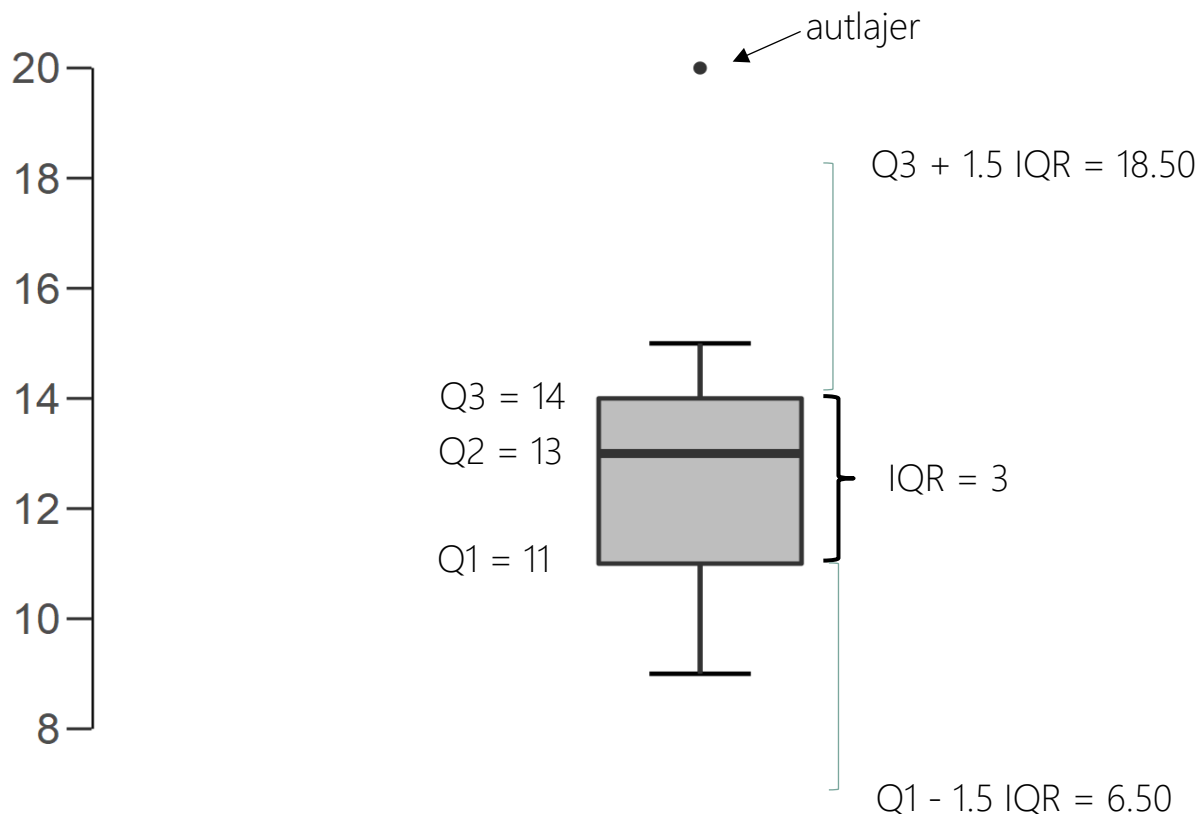
X	Statistik	Sa autlajerom	Bez autlajera
9			
11	Mod	11	11
11			
11	Medijana	13	13
11			
11	Interkvartilni raspon	3	3
12			
12	Medijansko apsolutno odstupanje	1.5	1
12			
13			
13			
13			
14			
14			
14			
14			
15			
15			
15			
20			



DETEKCIJA AUTLAJERA – METOD 1

Tukey (1977)

Vrednosti koji ste odstupaju više od 1.5 IQR od gornjeg ili donjeg kvartila





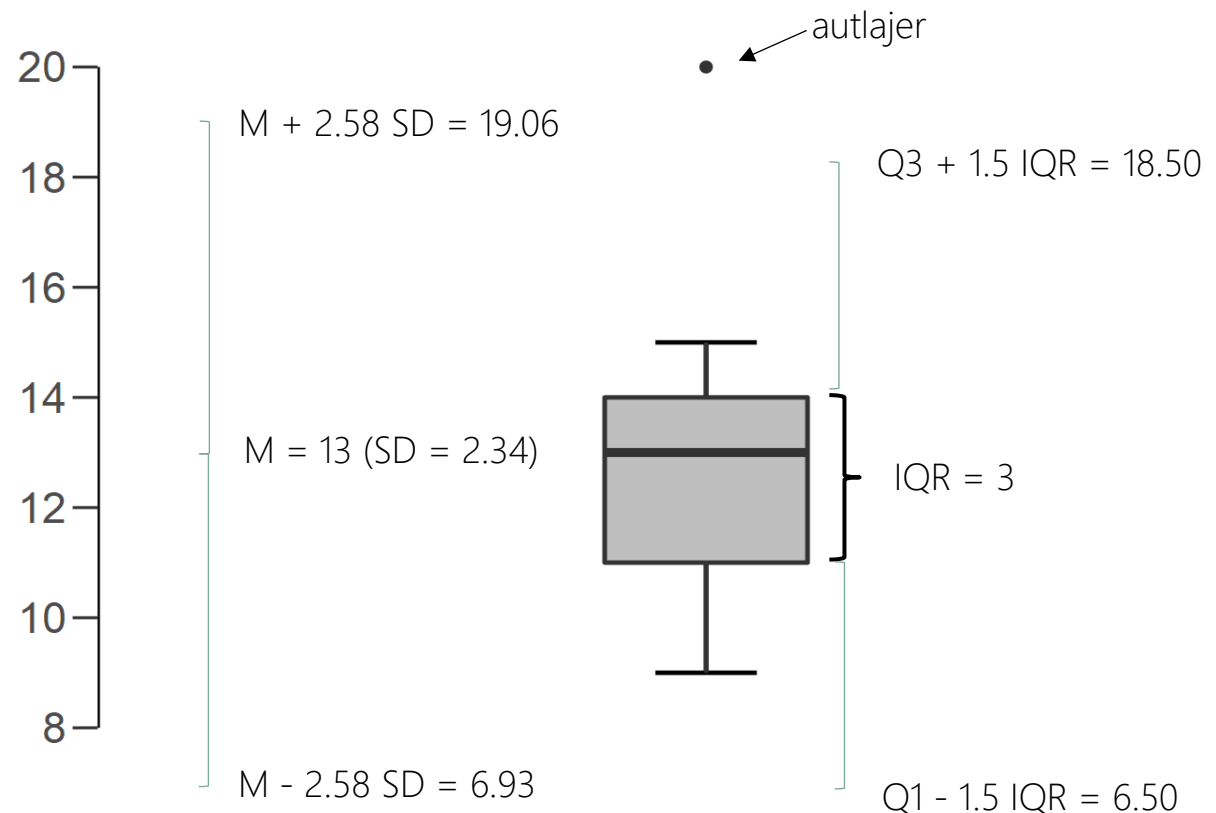
DETEKCIJA AUTLAJERA – METOD 2

Fidell & Tabachnik (2003)

Vrednosti koje odstupaju više od 3.29 standardnih devijacija od proseka (kada je $N > 100$)

2.58

$N < 100$

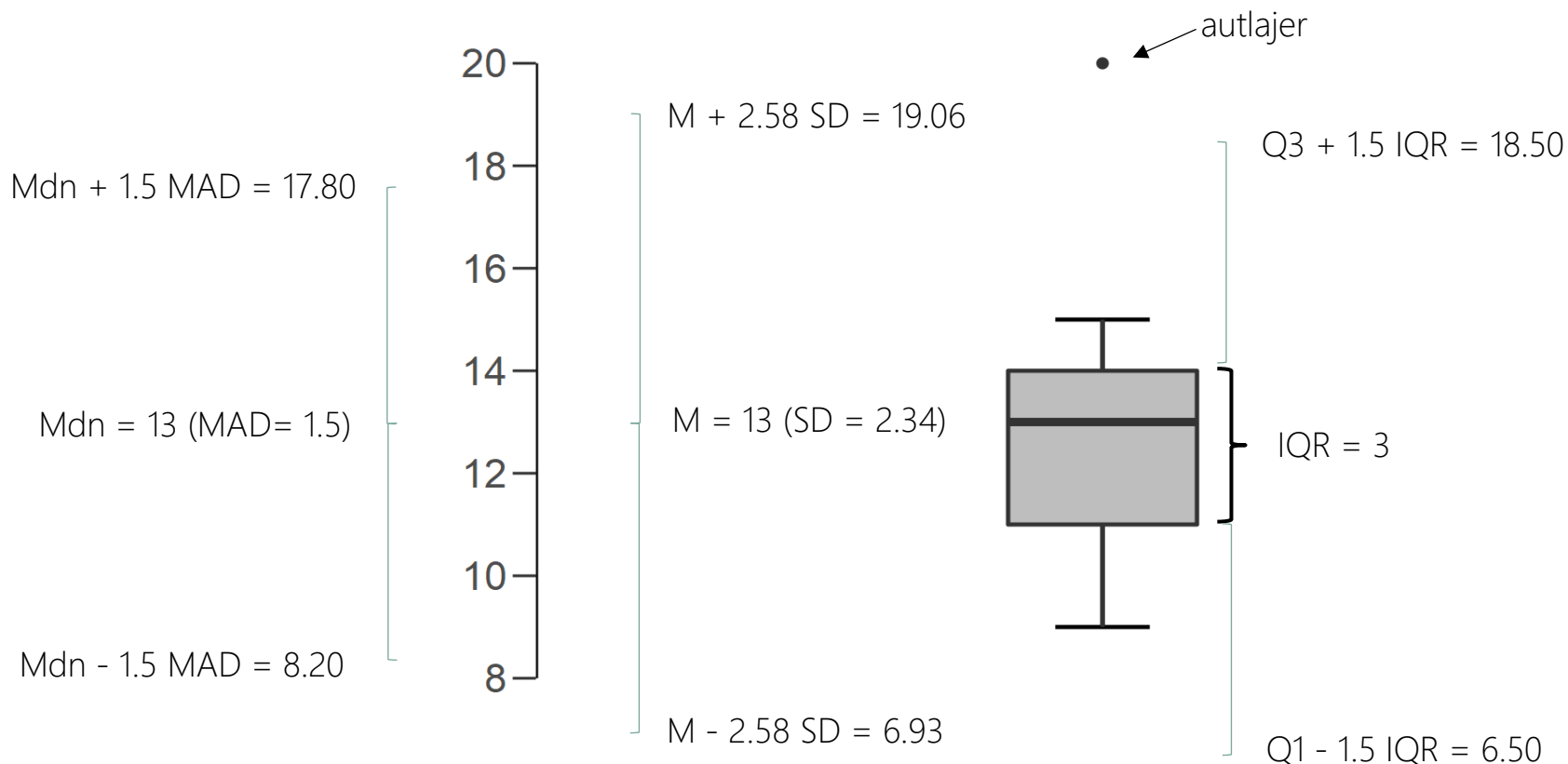




DETEKCIJA AUTLAJERA – METOD 3

Wilcox (2005)

Vrednosti koje odstupaju više od 3.20 medijanskih apsolutnih odstupanja od medijane





TRETIRANJE AUTLAJERA

~~20~~

Prva mogućnost – **odbacivanje**

Postriživanje (eng. trimming) ili odsecanje (eng. truncation)

Ne uzimamo u obzir autlajera u daljim analizama

18.5

~~20~~

Druga mogućnost – **vinzorizovanje**

Modifikacija vrednosti autlajera

Dodeljivanje najbliže prve prihvatljive vrednosti (npr. granice unutrašnje ograde)

20

Treća mogućnost – **zadržavanje**

Preporučena upotreba robusnih pokazatelja (postriženi prosek, MAD)

U inferencijalnoj statistici se preporučuje upotreba neparametrijskih testova