# Дескриптивна статистика

У овој лекцији ћеш учити о томе шта је дескриптивна статистика,

### Појмови које ћеш научити су:

* дескриптивна статистика
* средња вредност
* медијана
* мод
* кутијасти дијаграм
* стандардна девијација, варијанса
* нормална расподела

## Дескриптивна статистика

Као што је већ речено у претходној лекцији статистика је научни метод који се примењује са сврхом описивања, процене и анализе неке карактеристике друштва. Она представља грану математике која прикупља, анализира, тумачи, обрађује и приказује податаке чија је основна карактеристика варијабилност.

У претходној лекцији навели смо које су то две врсте статистике, а у овој ћеш учити о дескриптивној статистици.

Важна примена дескриптивне статистике је да сумира сакупљене податке на јасан и разумљив начин. Примењује се како би описали основна својства података који се анализирају. Заједно са једноставним графичким приказом, она даје и основну квантитативну анализу података.

## Средња вредност

Средња вредност или аритметичка средина представља суму свих података подељену са укупним бројем података. Означава се са .

Aко обележја имају фреквенције редом: , онда је:

На контролној вежби из математике ученици добили су 3 јединице, 7 двојки, 10 тројки, 7 четворки и 5 петица. Колика је средња оцена одељења на овој вежби?

У статистичком истраживању ово је једна од најважнијих нумеричких карактеристика испитиваног обележја, одмах након ње по значају налази се медијана.

## Медијана

Медијана je вредност средишњег податка која податке поређане по величини, почевши од најмањег до највећег, дели у два једнако бројна дела. Ако је број података непаран медијана је вредност средишњег податка, а ако је број података паран медијана представља средњу вриједност два средишња податка. Означава се са .

Медијана дели уређени низ података на два једнакобројна низа. Медијан прве половине низа је доњи (први), квартал и означава се са . Медијан друге половине низа назива се горњи (трећи), квартал и означавамо га са . Из овога произилази да медијану такође називамо и други квартал и такође је можемо означавати и са

У следећем низу одреди први, други и трећи квартал. Низ: 7, 9, 10, 10, 12, 14, 15, 15, 16, 17, 18.

(медијана)

Низ мањи од медијане је: 7, 9, 10, 10, 12. Први квартал представља медијану овог низа.

Низ већи од медијане је: 15, 15, 16, 17, 18. Трећи квартал представља медијану овог низа.

Из овог примера можеш да видиш да медијана дели низ података на четири једнака дела. У овом примеру ти делови би били: 7, 9 затим 10, 12 затим 15, 15 затим 17, 18.

Медијана није осетљива на екстремне вредности. Већина особина код људи је распоређена по нормалној расподели. Екстремне вредности се налазе на једном од два краја и захватају мали проценат резултата. Померање резултата ка крајевима дистрибуције не мења проценат резултата изнад и испод медијане. Због тога је медијана добра мера централне тенденције код изразито закривљених дистрибуција.

## Мод

Мод је вредност податка који се најчешће понавља (има нејвећу фреквенцију). Означава се са . Уколико се два податка појављују два пута у скупу оне је бимодалан. Фреквенција мода се одређује тако што након одређивања мода пребројимо колико пута се тај елемент понавља у низу.

Одредити мод низа, као и његову фреквенцију. Низ: 3, 5 ,5 ,7 ,11 ,11 ,11, 13, 15. Фреквенција мода је 3.

Одредити мод низа, као и његову фреквенцију. Низ: 3, 5 ,6 ,7 ,10 ,11, 13, 15, 17, 20.

Овај низ нема мод. Као што видиш и та могућност постоји!

Одредити мод низа, као и његову фреквенцију. Низ: 7, 9, 10, 10, 10, 12, 14, 15, 15, 15, 16 16, 17, 18.

Овај низ је бимодалан. Мод: 10 и 15, Фреквенција: 3.

## Кутијасти дијаграм

Уколико узмемо за пример аритметичку средину оцена ученика у једном одељењу која износи 3, долазимо до више могућих распореда оцена. Једна је да цело одељење има оцену 3 их предмета у питању, али такође случај може бити да половина има оцену 5, а половина оцену 1.

Ако скуп података желимо да представимо прецизније него што то чини аритметичка средина употребићемо кутијасти дијаграм. На њему се налази пет различитих података: минимална вредност података , максимална вредност података , први квартал , медијана и трећи квартал .

Кутијасти дијаграм

Кутијасти дијаграм

. Нацртати кутијасти дијаграм за дат низ. Низ: 8, 10, 12, 13, 15, 16, 18, 20, 25.

За почетак ћемо одредити неопходне податке.

Сада када смо одредили неопходне вредности можемо нацртати кутијасти дијаграм овог низа:

Кутијасти дијаграм

Кутијасти дијаграм

## Стандардна девијација

Расподела података око њихове аритметичке средине меримо одступањима појединиачних података од те средине. За укупну меру не можемо узети број тих одступања јер би се одступања с различитих страна међусобно поништавала (одступање оцене 1 од просека 3 износи . а оцине 5 износи , па је збир одступања ). Због овога ми узимамо квадрате тих одступања, како би резултат увек био позитиван (одступање оцене 1 од просека 3 износи . а оцине 5 износи , па је збир квадрата ових одступања )

Стандардну девијацију можемо да интерпретирамо као просечно одступање свих нумеричких података од њихове аритметичке средине. Самим тим закључујемо да што је она мања то су подаци ближи аритметичкој средини.

Стандардна девијација представља други корен из варијансе. Варијанса је мера расподеле података око аритметичке средине, дефинише се као просечно квадратно одступање од просека:

Из овога произилази да је стандардна девијација:

Одреди стандардну девијацију за следећи низ. Низ: 10, 12, 13, 15, 17, 18.

$ = $

$ =14 $

$ {} = $

$ {} = =2,77 $

Добијен резултат нам показује да дати подаци одступају од аритметичке средине за 2,77.

## Нормална расподела

Ако је податак о популацији (нпр. висина ученика у средњим школама) резултат многих међусобно независних утицаја (нпр. прехране, висине родитеља…), онда је хистограм релативних фреквенција тог податка увиек потпуно одређен његовом аритметичком средином $ $ и стандардном девијацијом $ {} $.

Има облик звона. Симетрична је у односу на праву $ x = $, kао код сваке симетричне криве аритметичка средина, модус и медијана су међусобно једнаки. X-оса је и са леве и са десне стране асимптота oве функције. $ {} $ јој одредује ширину распршења око $ $ на следећи начин:

Нормална расподела

Нормална расподела

### Питалице за крај:

1. Тони посљедњих седам дана мери број откуцаја срца након вежбања. Податке је поређао по величини: 98, 101, 101, 102, 103, 106, 110.

*Која је медијана записаних података:*

1. 101
2. 102
3. 103

*Tačno:* 2

*Која је аритметичка средина записаних података:*

1. 101
2. 102
3. 103

*Tačno:* 3

*Који је мод записаних података:*

1. 101
2. 102
3. 103

*Tačno:* 1

1. Кутијастим дијаграмом су приказани подаци. У складу са тим одговори на питања. Кутијасти дијаграм

*Шта представља број 33:*

1. $ $

*Tačno:* 3

*Који број од наведених представља :*

1. 42
2. 22
3. 27,5

*Tačno:* 2

*Који број од наведених представља $ Q\_1 $:*

1. 27,5
2. 42
3. 40

*Tačno:* 1

1. Ученици једнога разреда анкетирани су о броју деце у њиховој породици . Ово су њихови одговори: 1, 1, 2, 2, 2, 2. 2, 2, 2. 2. 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 4, 4, 6, 7.

*Шта бисте употребили како би приказали просечан број деце у породици:*

1. Мод
2. Медијану
3. Аритметичку средину

*Tačno:* 3

*Који je мод датог низа:*

1. 1
2. 2
3. 3

*Tačno:* 2

*Која је фреквенца мода низа:*

1. 6
2. 2
3. 16

*Tačno:* 3

1. Просечан број бодова које су постигли ученици на испитивању износи 142 бода, а стандардна девијација 7.

*Колико посто ученика има 135 - 149 бодова:*

1. 68,2 %
2. 95,4 %
3. 34,1 %

*Tačno:* 1

*У случају да је стандардна девијација била већа од 7 да ли би проценат ученика са 135 - 149 бодова био:*

1. Мањи
2. Већи
3. Исти

*Tačno:* 2

1. Ако је просечна висина ученица 165 cm, а њихова стандардна девијација 5 cm:

*Oдреди колико је посто ученица високо 160 - 170 cm:*

1. 68,2 %
2. 47,7 %
3. 34,1 %

*Tačno:* 1

*Oдреди колико је посто ученица високо 155 - 165 cm:*

1. 68,2 %
2. 47,7 %
3. 15,7 %

*Tačno:* 2

*Oдреди колико је посто ученица високо 170 - 180 cm:*

1. 15,7 %
2. 47,7 %
3. 34,1 %

*Tačno:* 1

*Колике су шансе да је ученица која припада oвој популацији виша од 180 cm:*

1. 3 %
2. 0 %
3. 0,1 %

*Tačno:* 3

*Колике су шансе да је ученица која припада oвој популацији нижа од 160 cm:*

1. 3 %
2. 20,4 %
3. 15,8 %

*Tačno:* 3