

Statistika përshkruese

Një udhëzues i afërt për ndryshoret, shpërndarjet dhe përmbledhjet numerike

Ratiomera Statistics

Përmbajtja

Përmbledhja e faqes së kursit	2
Shënimi dhe lista e kontrollit për vendime	2
Nivelet e matjes	2
Shpërndarjet e frekuencave	3
Prirja qendrore	3
Ndryshueshmëria	4
Transformimet lineare	4
Transformimi z	4
Zgjedhja e diagramit sipas nivelit të shkallës	4
Format e shpërndarjes	5
Grafikë që krijojnë pershtypje të rreme	5
Çfarë duhet të jesh në gjendje të bësh tani	5
Referencë e zgjeruar	6
Qëllimi dhe bazat	6
Idetë kryesore	6
Udhëzuesi i formulave	7
Si lexohet figura shpjeguese	9
Lista e kontrollit për interpretim	10
Si lidhet kjo temë me të tjerat	10

Përmbledhja e faqes së kursit

Statistika përshkruese e lidh matjen me analizën. Fillo duke përcaktuar se çfarë përfaqëson secili rast, ndryshore dhe vlerë, pastaj përdor nivelin e matjes për të zgjedhur frekuenca, përmbledhje numerike dhe grafikë të arsyetueshëm. Qendra, shpërhapja dhe forma duhen interpretuar së bashku, sepse asnjëra prej tyre nuk jep e vetme një përshkrim të plotë. Këto ide mbështesin probabilitetin, inferencën, korrelacionin, regresionin dhe çdo temë të mëvonshme të kësaj renditjeje mësimore.

Shënimi dhe lista e kontrollit për vendime

Tabela 1: Shënimi bazë që përdoret në statistikën përshkruese.

Simboli	Kuptimi	Pyetja që duhet të bësh
x_i	Vlera e vrojtimit i	Çfarë përfaqëson një vlerë e regjistruar?
n dhe N	Madhësia e kampionit dhe madhësia e popullatës	A po përshkruaj një kampion apo popullatën e plotë?
$\bar{x}$ dhe μ	Mesatarja e kampionit dhe mesatarja e popullatës	A përputhet simboli me objektivin?
s^2 dhe σ^2	Varianca e kampionit dhe varianca e popullatës	A përputhet emëruesi me objektivin?
s dhe σ	SD-ja e kampionit dhe SD-ja e popullatës	Cila është njësia origjinale e matjes?
Σ	Mblidh shprehjen e dhënë nëpër shtrirjen e indeksit të saj	Cili është termi i shumës dhe cilët janë kufijtë?

Përpara se të raportosh një rezultat, identifikoj rastet dhe ndryshoret, shqyrto vlerat që mungojnë ose janë të pamundura, përcakto nivelin e matjes, shqyrto shpërndarjen, zgjidh një qendër dhe shpërhapje të përputhshme dhe thuaj çfarë nuk mund të vërtetohet përmbledhjet për popullatën ose shkakësinë.

Nivelet e matjes

Pesë nivele të shkallës përcaktojnë se cilat veprime dhe përmbledhje janë të vlefshme. Çdo nivel i shton veti nivelit mbi të.

Tabela 2: Vetitë e matjes, transformimet e lejueshme dhe përmblendhjet kryesore.

Shkalla	Vetia përcaktuese	Rikodimi i lejueshëm	Përmblendhjet kryesore
Nominale	Kategori pa rend	Rietiketim një-për-një	Numërime, përpjesëtime, modë
Rendore	Kategori të renditura, largësi të pabarabarta	Rikodim rreptësisht rritës	Numërime, përpjesëtime, medianë, percentile, modë
Intervalore	Largësi të barabarta, zero e përcaktuar me marrëveshje	$y = a + bx, b > 0$	Mesatare, SD, diferenca, korrelacion
E raportit	Largësi të barabarta, zero e vërtetë	$y = bx, b > 0$	Përmblendhjet intervalore dhe raporte me kuptim
Absolute	Zero e vërtetë dhe njësi natyrore e pandryshueshme	$y = x$	Përmblendhjet e shkallës së raportit dhe numërimet

Shkallët intervalore, të raportit dhe absolute quhen së bashku shkallë **metrike**.

Shpërndarjet e frekuencave

Tabela 3: Referencë e përmblendhur për shpërndarjet e frekuencave.

Madhësia	Formula	Interpretimi
Frekuenca absolute	n_j	Numri i vrojtimeve në kategorinë j
Frekuenca relative	$f_j = n_j/n$	Përpjesëtimi në kategorinë j
Frekuenca relative kumulative	$F_j = \sum_{k=1}^j f_k$	Përpjesëtimi në kategorinë e renditur j ose më poshtë

Frekuencat absolute japin shumën n , ndërsa frekuencat relative japin shumën 1, përveç dallimeve nga rrumbullakimi. Frekuencat kumulative kërkojnë një ndryshore të renditur dhe nuk kanë kuptim për kategoritë nominale.

Prirja qendrore

Tri masa përshkruajnë vlerën tipike:

Tabela 4: Masat e prirjes qendrore dhe përdorimet e tyre.

Masa	Formula	Kur përdoret
Mesatarja	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	Shpërndarje simetrike; të dhëna metrike
Mediana	Vlera e mesit pas renditjes	Shpërndarje të anuara; rezistente ndaj vlerave të veçuara
Moda	Vlera ose kategoria më e shpeshtë	Çdo shkallë; mund të ketë barazim ose të mos jetë unike

Kur mesatarja është dukshëm më e lartë se mediana, ky model mund të sugjerojë asimetri djathtas, sidomos kur histogrami tregon edhe një bisht të gjatë djathtas. Vetëm renditja e dy statistikave nuk e provon formën e shpërndarjes.

Ndryshueshmëria

Katër masa përshkruajnë sa ndryshojnë individët:

Tabela 5: Masat e ndryshueshmërisë dhe përdorimet e tyre.

Masa	Formula	Kur përdoret
Amplituda	$x_{\max} - x_{\min}$	Pasqyrë e drejtpërdrejtë; e ndjeshme ndaj skajeve
IQR	$Q_3 - Q_1$	Shpërhapje rezistente; përdoret me medianën
Varianca	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	Baza për SD-në dhe statistika të tjera
SD	$s = \sqrt{s^2}$	Shkalla e devijimeve rreth mesatares; përdoret me mesataren

Vlerat $Q_1 - 1.5 \cdot \text{IQR}$ dhe $Q_3 + 1.5 \cdot \text{IQR}$ përcaktojnë kufirin e poshtëm dhe të sipërm. Mustaqet shtrihen deri te vlerat më të skajshme brenda këtyre kufijve. Vlerat përtej kufijve paraqiten veçmas si vlera të mundshme të veçuara.

Transformimet lineare

Transformimi linear $y_i = a + b \cdot x_i$ ka ndikime të parashikueshme në statistikat përmbledhëse:

Tabela 6: Ndikimet e transformimit linear në statistikat përshkruese.

Madhësia	Rregulli i transformimit	Ndikimi
Mesatarja	$\bar{y} = a + b\bar{x}$	Zhvendoset me a dhe shkallëzohet me b
Varianca	$s_y^2 = b^2 s_x^2$	Shkallëzohet me b^2 ; nuk ndikohet nga a
Devijimi standard	$s_y = b s_x$	Shkallëzohet me $ b $; nuk ndikohet nga a
Renditja dhe orientimi	Ruhen kur $b > 0$	Përmbysen kur $b < 0$

Mesatarja e transformuar ndjek formulën $\bar{y} = a + b\bar{x}$. Zhvendosja a e lëviz mesataren, por nuk ndikon në ndryshueshmëri, ndërsa $|b|$ i shkallëzon largësitë dhe devijimin standard. Kur b është negativ, shpërndarja pasqyrohet dhe renditja përmbysët.

Transformimi z

Për një kampion jokonstant me s të përcaktuar e jozero, transformimi $z_i = (x_i - \bar{x})/s$ prodhon mesatare të kampionit 0 dhe SD të kampionit 1.

Pikëzimi z të tregon sa devijime standarde mbi ose nën mesataren e referimit gjendet një vrojtim. Krahasimi i pikëzimeve z ndërmjet instrumenteve kërkon të njëjtin konstrukt, grupe të përshtatshme referimi dhe cilësi të krahasueshme të matjes.

Zgjedhja e diagramit sipas nivelit të shkallës

Tabela 7: Zgjedhja e diagramit sipas nivelit të matjes.

Niveli i shkallës	Diagrami i përshtatshëm
Nominal	Diagram me shtylla, diagram rrethor
Rendor	Diagram i renditur me shtylla
Metrik (intervalor, i raportit, absolut)	Histogram, diagram kuti-me-mustaqe

Në histograme, **sipërfaqja është në përpjesëtim me frekuencën**. Kur klasat kanë gjerësi të barabartë, lartësia dhe sipërfaqja e paraqesin drejt frekuencën. Kur gjerësitë nuk janë të barabarta, lartësia e dendësisë është f_j/w_j . Krahazo disa gjerësi të arsyeshme të klasave, sepse forma e dukshme mund të ndryshojë sipas grupimit.

Format e shpërndarjes

Tabela 8: Përmasa të pavarura që përdoren për të përshkruar formën e shpërndarjes.

Forma	Veçoria kryesore	Implikimi për interpretim
Simetrike	Bishta të barabartë në të dyja anët	Mesatarja dhe mediana janë pranë
E anuar djathtas	Bisht i gjatë drejt vlerave të larta	Mesatarja shpesh e tejkalon medianën; verifikoje grafikisht
E anuar majtas	Bisht i gjatë drejt vlerave të ulëta	Mesatarja shpesh bie nën medianë; verifikoje grafikisht
Unimodale	Një kulm	Një përqendrim kryesor, që përmbledhet sipas formës së tij
Bimodale	Dy kulme	Paraqiti dhe hetoji të dy kulmet; mos supozo nëngrupe
Kurtozë e lartë	Peshë më e madhe në devijimet e skajshme	Më e prirur për vrojtime të largëta
Kurtozë e ulët	Peshë më e vogël në devijimet e skajshme	Më pak e prirur për vrojtime të largëta

Simetria, modaliteti dhe kurtoza mund të bashkëjetojnë në kombinime të ndryshme. Vetëm lartësia e kulmit nuk e përcakton kurtozën.

Grafikë që krijojnë përshtypje të rreme

Grafikët mund të krijojnë përshtypje mashtruese përmes boshteve të shkurtuara, shkallëve të ngjeshura ose të zgjatura dhe klasave të papërshtatshme në histogram. Përqindjet mund të çojnë në përfundime të gabuara kur emëruesi, madhësia e kampionit, metoda e matjes ose të dhënat që mungojnë mbeten të fshehura. Si grafikët, ashtu edhe numrat kërkojnë kontekst dhe matje të vlefshme.

Çfarë duhet të jesh në gjendje të bësh tani

Tani duhet të jesh në gjendje të identifikosh rastet, ndryshoret, vlerat dhe nivelet e matjes; të lexosh shënimin që përdoret në mesatare, varianca dhe pikëzime z; të ndërtosh dhe interpretosh përmbledhje frekuencash; të zgjedhësh masa të përputhshme të qendrës dhe shpërhapjes; të krahasosh shpërndarje me tabela dhe grafikë; të parashikosh ndikimet e transformimeve lineare; të standardizosh vrojtime; dhe të dallosh zgjedhje paraqitjeje që mund ta shtrembërojnë një mesazh statistikor. Duhet gjithashtu të jesh në gjendje t'i shprehësh kufijtë e një përfundimi përshkrues pa e kthyer një model të vërtetuar në pohim për popullatën ose në pohim shkakor.

Referencë e zgjeruar

Qëllimi dhe bazat

Statistika përshkruese të jep një mënyrë të rregullt për ta kthyer një grup vrojtimesh në një përshkrim të kuptueshëm të asaj që është vrojtuar. Fillor duke përcaktuar **rastet**, pra njerëzit ose njësitë që përfaqësohen nga rreshtat e një grupi të dhënash, dhe **ndryshoret**, pra karakteristikat e regjistruara në kolonat e tij. Një vlerë është një rezultat i regjistruar për një ndryshore të një rast. Para se të llogaritësh, pyet çfarë do të thotë ndryshorja, si është matur dhe cilat vlera mund të marrë. Kjo e parandalon një llogaritje matematikisht të saktë të kthehet në një përshkrim çorientues.

Niveli i matjes udhëzon çfarë mund të bësh në mënyrë të arsyeshme me një ndryshore. Një ndryshore nominale i ndan rastet në kategori pa rend. Një ndryshore rendore ka kategori të renditura, por largësitë mes kategorive fqinje nuk dihet nëse janë të barabarta. Një ndryshore intervalore ka largësi të barabarta me kuptim, por nuk ka një zero absolute me kuptim. Një ndryshore e raportit ka largësi të barabarta dhe një zero me kuptim, prandaj mund të interpretohen edhe raportet. Niveli i matjes është veti e mënyrës si përkufizohet dhe matet ndryshorja. Ai nuk varet nga pamja që marrin vlerat e saj në një grup të vetëm të dhënash.

Shkalla	Çfarë tregojnë vlerat	Përmbledhje të përshtatshme fillestare
Nominale	Nëse rastet i përkasin së njëjtës kategori apo kategorive të ndryshme	Frekuencat, përpjesëtimet, moda
Rendore	Përkatesinë në kategori dhe rendin	Frekuencat, përpjesëtimet, mediana, kuantilet
Interval/raport	Rendin dhe largësinë numerike me kuptim	Mesatarja, mediana, varianca, devijimi standard, kuantilet

Disa nga grupet e të dhënave që përdoren në këtë rrjedhë mësimore janë **të dhëna të simuluar**, domethënë vlera të krijuara nga kompjuteri sipas rregullave të deklaruara dhe jo matje të mbledhura nga njerëz realë. **Simulimi** është procesi që i krijon këto vlera. Kompjuteri përdor një **gjenerator të numrave të rastësishëm**, pra një algoritëm të ndërtuar për të prodhuar vlera që sillen si rezultate rastësie. **Numri nisës** është vlera fillestare që i jepet atij algoritmi. Kur ripërdoret i njëjti numër nisës me të njëjtat udhëzime, rikrijohet i njëjti grup i të dhënave. Kjo e bën shembullin mësimor të riprodhueshëm: ti dhe një person tjetër që mëson mund të shqyrtoni të njëjtat vrojtme dhe të merrni të njëjtat rezultate. Simulimi e mbështet të nxënit, por nuk i shndërron vlerat e krijuara në evidencë për një popullatë reale.

Idetë kryesore

Një shpërndarje përshkruan si shtrihen vlerat e një ndryshoreje në diapazonin e tyre të mundshëm. Për një ndryshore kategorike, fillor me një tabelë frekuencash. Frekuenca absolute është numri i rasteve në një kategori. Frekuenca relative është ky numër i pjesëtuar me numrin e përgjithshëm të rasteve të vlefshme. Frekuencat relative mund të paraqiten si përpjesëtime ose përqindje. Kontrolllo gjithmonë nëse vlerat që mungojnë janë përjashtuar nga emëruesi, sepse një përqindje ka kuptim vetëm kur dihet totali i saj referues.

Për një ndryshore numerike, përshkruaj së bashku katër veçori: qendrën, ndryshueshmërinë, formën dhe vrojtmet e pazakonta. Mesatarja përdor çdo vlerë dhe përfaqëson pikën e baraspeshës së shpërndarjes. Mediana është vlera e mesit pasi vlerat renditen dhe i ndan të dhënat në dy gjysma. Moda është vlera ose kategoria më e shpeshtë. Amplituda shtrihet nga vlera më e vogël te vlera më

e madhe. Diapazoni ndërkuartilor përfshin gjysmën e mesme të vrojtmeve të renditura. Varianca dhe devijimi standard përmbledhin sa larg prirën të jenë vlerat nga mesatarja.

Pyetja	Evidencë e dobishme	Zakon i dobishëm leximi
Ku gjendet qendra e shpërndarjes?	Mesatarja, mediana dhe ndonjëherë moda	Krahaso mesataren me medianën në vend që të raportosh vetëm një numër
Sa ndryshojnë vrojtmet?	Amplituda, diapazoni ndërkuartilor, varianca, devijimi standard	Trego njësitë dhe vëzhgo vlerat e pazakonta
Çfarë forme krijojnë vlerat?	Histogrami, diagrami kuti-me-mustaqe, frekuencat, asimetria	Kërko simetri, anim, boshllëqe, grumbullime dhe më shumë se një kulm
A ka vlera befasuese?	Vlerat e papërpunuara, grafiku, kontrollet e të dhënave, pikëzimet e standardizuara	Heto para se të vendosësh nëse një vlerë është gabim

Forma ndikon në interpretim. Në një shpërndarje afërsisht simetrike, mesatarja dhe mediana shpesh janë të ngjashme. Një bisht i gjatë djathtas priret ta tërheqë mesataren lart, ndërsa një bisht i gjatë majtas priret ta tërheqë poshtë. **Modaliteti** përshkruan numrin dhe modelin e kulmeve ose grumbullimeve kryesore që dallohen qartë. Një shpërndarje mund të jetë unimodale, me një kulm kryesor, ose multimodale, me më shumë se një grumbullim. **Kurtoza** përshkruan sa lehtë shfaqen vlerat larg në bishta, krahasuar me një shpërndarje referuese simetrike në formë kambane me të njëjtën shpërhapje të përgjithshme. Vetëm lartësia e kulmit nuk e përcakton kurtozën. Një numër i vetëm si mesatarja nuk mund t'i tregojë këto veçori, prandaj grafiku dhe përmbledhjet numerike duhen lexuar së bashku.

Një vrojtim i pazakontë nuk është automatikisht gabim. Mund të jetë rast i vlefshëm por i rrallë, gabim kodimi, problem matjeje ose shenjë se janë bashkuar grupe të ndryshme. Kontrolllo përkufizimin fillestar dhe procesin e regjistrimit para se të përjashtosh diçka. Nëse një vendim analitik ndryshon pasi hiqet një vrojtim, raportoje këtë ndjeshmëri dhe mos e fshih.

Udhëzuesi i formulave

Për kategorinë ose intervalin j , le të jetë n_j frekuenca e tij absolute dhe n numri i vrojtmeve të vlefshme. Frekuenca relative e tij është

$$f_j = \frac{n_j}{n}.$$

Për kategori të renditura ose intervale numerike, frekuenca relative kumulative deri te kategoria j është

$$F_j = \sum_{h=1}^j f_h.$$

Frekuencat absolute duhet të japin gjithsej n , frekuencat relative duhet të japin 1 me përjashtim të ndryshimeve nga rrumbullakimi, ndërsa frekuenca relative kumulative e fundit duhet të jetë 1. Frekuencat kumulative kanë kuptim vetëm kur kategoritë kanë një rend që mund të arsyetohet.

Le të jenë $x_1, x_2, \dots, x_n$ vlerat e vrojtura të një ndryshoreje numerike. Mesatarja e kampionit i mbledh të gjitha vlerat dhe e pjesëton shumën me numrin e vrojtmeve:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Shenja $\sum$ do të thotë «mbledh vlerat e treguara». Indeksi i identifikon nga një vrojtje, duke filluar nga vrojtja e parë e deri te vrojtja n . Devijimi $x_i - \bar{x}$ tregon sa larg mbi ose nën mesatare gjendet një vlerë. Kur mbledhen, devijimet pozitive dhe negative e anulojnë njëra-tjetrën. Prandaj, për variancën ato ngrihen fillimisht në katror. Varianca e kampionit përdor $n - 1$ në emërues:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Varianca përkatëse e popullatës përdor mesataren e popullatës μ dhe pjesëton me madhësinë e popullatës N :

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2.$$

Mbaji të ndara dy madhësitë që synohen. Emëruesi $n - 1$ i përket variancës së korrigjuar të kampionit, e cila përdoret për të vlerësuar ndryshueshmërinë e popullatës. Pjesëtimi me N përshkruan vetë vlerat e popullatës së plotë.

Meqë varianca shprehet në njësi të ngritura në katror, marrim rrënjën e saj katrore për t'u kthyer te njësia fillestare e matjes. Prandaj devijimi standard i kampionit është:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Një pikëzim i standardizuar e shpreh një vlerë në njësi të devijimit standard. Ai zbret mesataren e kampionit dhe pjesëton me devijimin standard të kampionit:

$$z_i = \frac{x_i - \bar{x}}{s}$$

Një z_i pozitiv e vendos vlerën mbi mesatare, një z_i negativ e vendos nën mesatare, ndërsa madhësia e tij tregon largësinë në devijime standarde. Standardizimi ndryshon njësinë dhe pikën e referimit, por nuk ndryshon rendin ose formën e vrojtimeve.

Mediana dhe kuantilet fillojnë nga vrojtjet e renditura. Kuartili i parë Q_1 shënon çerekun e poshtëm, mediana Q_2 shënon pikën e mesit dhe kuartili i tretë Q_3 shënon tre të katërtat e poshtme. Amplituda dhe diapazoni ndërkuartil janë

$$\text{amplituda} = x_{\max} - x_{\min}, \quad IQR = Q_3 - Q_1.$$

Konventat për kuantilet e kampionit mund të ndërfusin vlera në mënyra të ndryshme, prandaj programet mund të raportojnë kuartile paksa të ndryshme për një grup të vogël të dhënash. Një kontroll i zakonshëm në diagramin kuti-me-mustaqe përdor kufijtë e brendshëm

$$Q_1 - 1.5(IQR) \quad \text{dhe} \quad Q_3 + 1.5(IQR).$$

Vlerat përtej një kufiri janë vlera të mundshme të veçuara që duhen shqyrtuar, jo gabime që duhen fshirë automatikisht. Mustaqet ndalen te vlerat më skajore të vrojtuar që ende gjenden brenda kufijve. Ato nuk përfundojnë domosdoshmërisht pikërisht te vlerat e kufijve.

Për një transformim linear $Y = a + bX$, qendra dhe shpërhapja ndryshojnë sipas

$$\bar{y} = a + b\bar{x}, \quad s_y^2 = b^2 s_x^2, \quad s_y = |b| s_x.$$

Zhvendosja a e ndryshon vendndodhjen, por jo shpërhapjen. Shumëzuesi b i ndryshon largësitë me $|b|$, prandaj varianca ndryshon me b^2 . Standardizimi është rasti i veçantë që zbret mesataren dhe pjesëton me devijimin standard.

Lartësia e histogramit kërkon një kontroll të fundit. Nëse intervali j ka frekuencë relative f_j dhe gjerësi w_j , lartësia e dendësisë është

$$h_j = \frac{f_j}{w_j}.$$

Atëherë sipërfaqja e shtyllës është $h_j w_j = f_j$. Kur intervalet kanë gjerësi të barabartë, lartësia mund ta pasqyrojë drejtpërdrejt frekuencën. Kur gjerësitë ndryshojnë, duhen përdorur lartësitë e dendësisë, në mënyrë që sipërfaqja dhe jo vetëm lartësia të vazhdojë ta përfaqësojë frekuencën.

Si lexohet figura shpjeguese

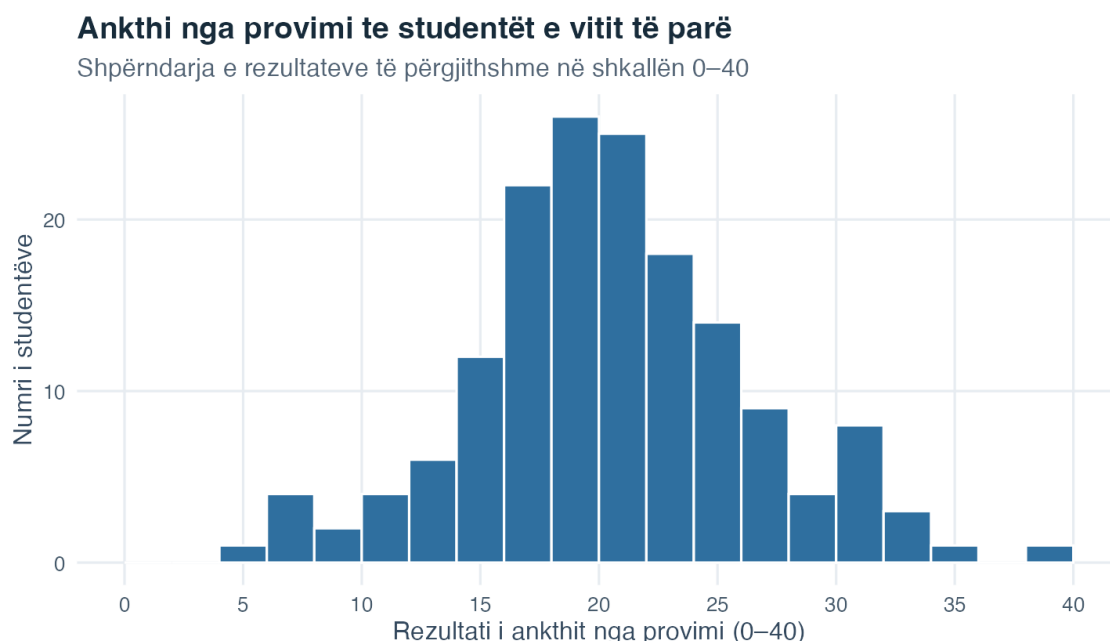


Figura 1: Histogram i pikëzimeve të simuluar të ankthit nga provimi prej zeros deri në dyzet, me shtyllat më të larta pranë qendrës dhe bishta më të rrallë në të dy skajet.

Lexo fillimisht boshtin horizontal. Ai jep pikëzimet e ankthit nga provimi në një shkallë nga 0 deri në 40. Boshti vertikal jep numrin e studentëve, prandaj lartësia e secilës shtyllë është një frekuencë. Shtyllat rreth pikëzimeve 18 deri në 22 janë më të lartat, gjë që e vendos grumbullimin kryesor pranë qendrës së shkallës. Në skajin e poshtëm dhe të sipërm ka më pak vrojtime. Një shtyllë pranë 40 tregon se ka të paktën një vlerë të lartë, por vetëm grafiku nuk tregon nëse ajo vlerë është e gabuar. Para se ta gjykosh, duhet të kthehesh te përkufizimi i të dhënave dhe procesi i regjistrimit.

Shtyllat prekin njëra-tjetrën sepse një shkallë numerike vazhdon përmes intervaleve fqinje. Gjerësia e tyre ka rëndësi: ndryshimi i kufijve të intervaleve mund t'i bëjë të njëjtat vrojtime të duken më shumë ose më pak të hollësishme. Prandaj figura duhet lexuar si pamje e një shpërndarjeje dhe jo si grup kategorish të ndara. Ajo mbështet pohime për qendrën, shpërhapjen, formën dhe vrojtimit e pazakonta. Nuk tregon një shkak të ankthit, nuk krahason një popullatë me një tjetër dhe nuk vërteton se modeli i simuluar ndodh te studentët realë.

Lista e kontrollit për interpretim

Fillo çdo përshkrim me rastet, ndryshoren, nivelin e saj të matjes dhe diapazonin e vlefshëm. Trego numrin e vrojtimeve të vlefshme dhe atyre që mungojnë. Për një ndryshore kategorike, raporto frekuencat bashkë me emëruesin dhe përpjesëtimet. Për një ndryshore numerike, bashko një grafik me masa të qendrës dhe ndryshueshmërisë. Përdor mesataren dhe devijimin standard kur interpretimi i tyre i përshtatet shpërndarjes. Shqyrto edhe medianën dhe diapazonin ndërkuartilor kur kanë rëndësi animi ose vrojtimit e pazakonta.

Mbaje të dukshme njësinë. Një devijim standard prej pesë pikësh do të thotë diçka tjetër nga pesë orë. Mos e përshkruaj një grup si homogjen ose të ndryshueshëm pa një referencë që e bën krahasimin me kuptim. Kontrolllo nëse një tabelë e rumbullakosur ende jep totalin e pritur. Shënoji të dhënat e simuluar si të simuluar. Në fund, ndaj përshkrimin nga shpjegimi: një model në vlerat e vrojtuar tregon çfarë përmban grupi i të dhënave, ndërsa një shpjegim shkakor kërkon dizajn kërkimor dhe arsyetim përtej statistikës përshkruese.

Si lidhet kjo temë me të tjerat

Statistika përshkruese siguron gjuhën që përdoret në tërë kursin. Probabiliteti shton rregulla për arsyetimin mbi rezultate të pasigurta. Inferenca statistikore përdor pastaj një kampion dhe ndryshueshmërinë e tij për të bërë pohime të kujdesshme për një popullatë. Kovarianca dhe korrelacioni pyesin si ndryshojnë së bashku dy ndryshore. Regresioni e shpreh një ndryshore rezultati si funksion të një ose më shumë ndryshoreve parashikuese, ndërsa korrelacioni i pjesshëm studion një lidhje pas përshtatjes lineare. Analiza e variancës krahason mesataret e grupeve duke e ndarë ndryshueshmërinë e përgjithshme në përbërës me kuptim.

Zakoni qendror mbetet i pandryshuar: kupto ndryshoret, shqyrto shpërndarjen, llogarit një përmbledhje të përshtatshme dhe interpretoje në kontekst. Metodat e mëvonshme shtojnë pasiguri dhe modele, por nuk e heqin kurrë nevojën për një përshkrim të besueshëm të të dhënave që hynë në analizë.

ID-ja e dokumentit: topic-01-descriptive-statistics-summary-sq

Shënime 1 nga 3

Statistika përshkruese | Tema 1

Shënime 2 nga 3

Statistika përshkruese | Tema 1

Shënime 3 nga 3

Statistika përshkruese | Tema 1