1.) POPOLN SISTEM DOGODKOV
Množico dogodkov $S=\{A_1,A_2,\dots A_n\}$ imenujemo popoln sistem dogodkov, če se v vsaki ponovitvi poskusa zgodi natanko eden od dogodkov iz množice S.To pomeni, da so vsi mogoči $A_i \neq N$ paroma nezdružljivi $A_i \cap A_j = N$ $i \neq j$ in je njihova vsota gotov dogodek $A_1 \cup A_2 \cup \dots \cup A_n = G$
Primer: Popoln sistem dogodkov pri metu kocke sestavljajo npr. elementarni dogodki ali pa tudi dva dogodka: dogodek, da vržem sodo št. pik in dogodek, da vržem liho št. pik.

2.) OSNOVNI POJMI VERJETNOSTNEGA RAČUNA (poskus, dogodek, def. verjetnosti)
Verjetnostni račun obravnava zakonitosti, ki se pokažejo v velikih množicah enakih ali vsaj zelo podobnih pojavov. Predmet verjetnostnega računa je torej empirične narave, njegovi osnovni pojmi pa so prevzeti iz izkušnje. POSKUS – je realizacija neke množice skupaj nastopajočih dejstev. Poskus je torej vsako dejanje, ki ga opravimo v natanko določenih pogojih. Rezultatu poskusa rečemo IZID množici vseh izidov pa PROSTOR IZIDOV. *Primer: met igralne kocke,*DOGODEK – je pojav, ki
ne spada v množico skupaj nastopajočih dejstev in se lahko v posameznem poskusu zgodi ali pa ne.
Primer: pri poskusu meta igralne kocke pade 6 pik, pri poskusu izbire karte iz kupa 20 kart izvlo rdečo , …
GOTOV DOGODEK – G: se zgodi ob vsaki ponovitvi poskusa.
Primer: vržemo 1,2,3,4,5,6 pik pri metu igralne kocke NEMOGOČ DOGODEK – N: se nikoli ne zgodi.
Primer: vržemo 7 pik pri metu igralne kocke SLUČAJEN DOGODEK – se včasih zgodi, včasih pa ne.
Primer: vržemo 6 pik pri metu igralne kocke A c B – podmnožica, dogodek A je način dogodka B (vsakič ko se zgodi A, se zgodi tudi B). Če je A c B in B c A => A=B *Primer: dogodek A ={ pade 6 pik} je način dogodka B = {pade sodo št. pik}* $A \cap B$ – presek, produkt dog. A in B (če se zgodita oba dog.hkrati)
Primer: produkt dogodka A ={vržemo sodo št. pik} in B ={vržemo liho št. pik} je nemogoč dogodek
A U B – unija, vsota dogodkov A in B (če se zgodi vsaj eden od dogodkov A in B).
Primer: vsota dogodka A ={vržemo sodo št. pik} in B ={vržemo liho št. pik} je gotov dogodek
A^c – komplement, nasprotni dogodek (negacija dogodka A)
Primer: nasproten dogodek, da vržemo sodo št. pik je, da vržemo liho št. pik
PRAVILA: A U N = A; A U G = G; A U B = B U A; A U (B U C) = (A U B) U C; $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$; $(A \cup B)^c = A^c \cap B^c$; $(A^c)^c = A$; $N^c = G$; $A \cap N = N$; $A \cap G = A$
NEZDRUŽLJIVA DOGODKA – A in B sta nezdružljiva, če se ne moreta zgoditi hkrati
Primer: dogodka A ={vržemo sodo št. pik} in B ={vržemo liho št. pik} sta nezdružljiva
Poljuben dogodek in njegov nasprotni dogodek sta vedno nezdružljiva. Njun produkt je nemogoč dogodek, njuna vsota pa gotov dogodek. SESTAVLJEN DOGODEK – je vsota nezdružljivih in mogočih dogodkov. Dogodek, ki ni sestavljen je ELEMENTAREN. *Primer: pri metu kocke je 6 el. dog(1 pika, 2 piki,…). Dogodek, da pade sodo št. pik pa je sestavljen dogodek.*
VERJETNOST – denimo, da smo n-krat ponovili dan poskus in da se je k-krat zgodil dogodek A. Ponovitve poskusa, v katerih se zgodi A, imenujemo ugodne za dogodek A, število f(A) = k/n pa je rel. frekv. dog.A. Če poskus X dolgo ponavljamo, se relativna frekvenca slučajnega dogodka ustali.
STATISTIČNA DEF. VERJETNOSTI – verjetnost dogodka A v danem poskusu je število P(A), pri katerem
se navadno ustali relativna frekvenca dogodka A v velikem številu ponovitev tega poskusa.
LASTNOSTI VERJETNOSTI: $P(A) \geq 0$ (relativna frekvenca je vedno nenegativna),
P(G) = 1, P(A U B) = P(A) + P(B) (za A in B nezdružljiva dogodka),
$P(A \cup B) = P(A) + P(B) – P(A \cap B)$ (za A in B združljiva dogodka), $P(A^c) = 1 – P(A)$
KLASIČNA DEF. VERJETNOSTI – Vzemimo da so dogodki iz popolnega sistema dogodkov $\{E_1, E_2,\dots,E_s\}$ enako verjetni: $P(E_1) = P(E_2)=\dots= P(E_s) = p$ Tedaj je verj. 1 od dogodkov $P(E_i) = 1/s$ za $i = 1,\dots,s$

Če je nek dogodek A sest. iz r dog. iz tega popolnega sistema dog., potem je njegova verjetnost P(A) = r/s

3.) POGOJNA VERJETNOST Zanima nas verjetnost dogodka A, če vemo da se je zgodil dogodek B (P(B)>0). Taki verjetnosti pravimo pogojna verjetnost dogodka A glede na dogodek B. Označimo jo s P(A|B), izračunamo pa kot $P(A|B) = P(A \cap B) / P(B)$ kjer je pogoj P(B)>0 in velja $0 \leq P(A|B) \leq 1$. Dogodka A in B sta neodvisna, če nam dogodek B 'nič ne pove' o dogodku A. Torej A in B sta neodvisna ntk velja P(A|B)=P(A) oz. $P(A \cap B)=P(A)P(B)$ Utemeljitev P(A|B) s pomočjo statistične def. Verjetnosti: Denimo, da smo n-krat ponovili poskus X in da se je ob tem k_B-krat zgodil dogodek B. To pomeni, da smo v n ponovitvah poskusa X napravili k_B-krat poskus X'. Dogodek A se je zgodil ob poskusu X' le, če se je zgodil tudi B => $A \cap B$. Denimo, da se je dogodek $A \cap B$ ob ponovitvi poskusa zgodila $k_{A \cap B}$-krat. In je rel.fr. dog A v opravljenih ponovitvah posk X': $f'(A) = f(A|B) = k_{A \cap B} / k_B = k_{A \cap B}|n / k_B|n = f(A \cap B) / f(B)$

4.) RELEJNI (DVOFAZNI) POSKUSI, FORMULA ZA POPOLNO VERJETNOST IN BAYESOV OBRAZEC

Dogodki $H_1, H_2, \ldots, H_n$ tvorijo popoln sistem dogodkov, če so paroma nezdružljivi $H_i \cap H_j = N$ in je njihova vsota gotov dogodek $H_1 \cup H_2 \cup \ldots \cup H_n = G$. Denimo, da naš poskus poteka v dveh fazah in da v prvi fazi nastopi natanko en dogodek iz popolnega sistema dogodkov, od njega pa je odvisen nek dogodek v drugi fazi. Denimo, da se je v drugi fazi zgodil dogodek A. Poznamo H_i, zanima pa nas $P(A|H_i)$. Formula za **popolno verjetnost**: $P(A) = \sum_{i=1}^{n} P(A|H_i)P(H_i)$ Poznamo H_i, zanima pa nas $P(H_i|A)$ torej verj., da se je A zgodil s točno določenim dogodkom iz prve faze.

Bayesov obrazec: $P(H_i|A) = P(H_i)\, P(A|H_i) / \sum_{i=1}^{n} P(A|H_i)P(H_i)$

5.) DISKRETNA SLUČAJNA SPREMENLJIVKA (ENAKOMERNA, BINOMSKA)

Denimo, da imamo poskus, katerega izidi so števila (npr. pri metu kocke števila). Poskusu je prirejena neka količina, ki ima različne vrednosti. Torej je spremenljivka. Katero od mogočih vrednosti zavzame v določeni ponovitvi poskusa, je odvisno od slučaja, zato ji rečemo **slučajna spremenljivka**.

Da je sluč. spr. **z**nana moramo poznati njeno zalogo vrednosti in verjetnost vsake izmed možnih vrednosti ali intervala vrednosti (predpis, ki določa te verj.imenujemo porazdelitveni zakon). $(X = x_i)$ – dogodek, da sluč. spr. X zavzame vrednost x_i. $F(x) = P(X < x)$ je porazdelitvena funkcija. **Diskretna sluč. spr.** – zaloga vrednosti je končna množica. $\{x_1, .., x_m\}$. Dogodki $X = x_k$ za $k = 1,2,\ldots,n$ tvorijo popoln sistem dogodkov. Verjetnost posameznega dogodka $P(X = x_i) = p_i$ Vsota verjetnosti vseh dogodkov je enaka 1 $p_1+p_2+\ldots+p_m = 1$ Verjetnostna shema prikazuje spremenljivko s tabelo tako, da so zgoraj zapisane vrednosti x_i, pod njimi pa pripadajoče verjetnosti: X: $\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}$. Porazdelitvena funkcija je v tem primeru $F(x_k) = P(X< x_k) = \sum_{1=1}^{k-1} p_i$

<u>Enakom. disk. porazdelitev</u> – diskretna sluč. spr. se porazdeljuje enakomerno, če so vse njene vred.i enako verjetne.

Primer: št. pik pri metu poštene igralne kocke. Vsaka pika ima verjetnost 1/6.

<u>Binomska porazdelitev</u> – zaloga vrednosti $\{0,1,2,\ldots,n\}$ in verjetnosti, ki jih izračunamo po Bernoullijem obrazcu

$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ za $k=0,1,2,\ldots,n$ Binomska porazdelitev je natanko določena z dvema podatkoma: n in p. Sluč. spr. X se porazdeljuje binomsko s parametroma n in p: X: b(n,p)

6.) BERNOULLIJEVO ZAPOREDJE NEODVISNIH POSKUSOV

Zaporedje neodvisnih poskusov se imenuje Bernoullijevo zaporedje, če se mora zgoditi v vsakem poskusu iz
zaporedja neodvisnih poskusov le dogodek A z verjetnostjo $P(A) = p$ ali dogodek A^c z verjetnostjo $P(A^c) = 1-P(A)=1-p=q$. *Primer: met kocke, kjer pri vsaki ponovitvi poskusa pade 6 z verje. P(A)=1/6 ali pa*
ne pade 6 z verjetnostjo $P(A^c)=q=5/6$. **Bernoullijev obrazec:** Oglejmo si najprej primer, ko v prvih k metih
pade cifra, v naslednjih n−k pa grb. Tak dogodek se zgodi z verjetnostjo $p^k(1-p)^{n-k}$.
Zdaj pa si oglejmo, koliko "podobnih dogodkov" imamo. Če točno vemo, na
katerih mestih je padla cifra, je seveda jasno, da je na preostalih padel grb. Zanima nas torej, na koliko načinov lahko določimo k mest izmed n možnih. Iz kombinatorike vemo, da je takih razporedb $\binom{n}{k}$.
Iskana verjetnost je torej $\binom{n}{k}p^k(1-p)^{n-k}$.

7.) ZVEZNA SLUČ SPR (PORAZD FUNKCIJA, VERJETN GOSTOTA)

Zvezna sluč. spr. – lahko zavzame vsako realno št. znotraj določenega intervala $a \le X \le b$. Verjetnost, da sluč.
spr. zavzame vrednost manjšo od neke vrednosti x, je: $F(x) = P(X<x) = \int_a^x p(x)dx$, kjer p(x) imenujemo **verjetnostna gostota**. To verjetnost si lahko predstavimo tudi grafično v koordinatnem sistemu, kjer na x-os nanašamo vrednosti slučajne spremenljivke, na y-os pa verjetnostno gostoto p(x). Verjetnost je tedaj predstavljena kot ploščina pod krivuljo, ki jo določa p(x). Velja: $\int_a^x p(x)dx=1$.
Porazdelitvena funkcija $F(x) = P(X < x)$ nam pove kakšna je verj., da je sluč. spr. X manjša od neke vred x e **R**.

8.) NORMALNA PORAZDELITEV (IN STANDARDIZIRANA NORMALNA PORAZDELITEV)

Sluč. spr. X, ki ima gostoto $p_X(x) = (1/\sigma\sqrt{2\Pi})e^{-(x-\mu)^2/2\sigma^2}$ je porazdeljena normalno s parametroma σ in μ. Označimo $X\sim N(\mu,\sigma)$. μ je mat.up. in σ st.i odklon normalno porazdeljene sluč. spr.
Če to sluč. spr. standardiziramo $Z = (X – \mu) / \sigma$ je Z še vedno normalno porazdeljena s parametroma Z: $N\sim(0,1)$. Tedaj pravimo, da je spr porazdeljena **stan.norm**. Gostota se tedaj poenostavi $v = (1/\sqrt{2\Pi})e^{-x^2/2}$

V splošnem nas zanimajo verj. dog, da zavzame sluč. spr. X vrednosti v intervalu $[x_1,x_2]$: $P(x_1<X< x_2) = \int_{x_1}^{x_2}p(x)dx$
Velja: $P(x_1<X< x_2) = P((x_1-\mu)/\sigma < (X-\mu)/\sigma < ((x_2-\mu)/\sigma) = P(z_1<Z<z_2)$
Standardizirano normalno porazdeljena sluč. spr. Z ima $E(Z)=0$ in $D(Z)=1$.
Za $Z\sim N(0,1)$ so verjetnosti $P(0\le Z\le z_1)=\Phi(z_1)$ izračunane in podane v tabeli: $\Phi(-z_1) = -\Phi(z_1)$
$P(Z>z_1) = 0{,}5-\Phi(z_1)$ $P(z_1<Z< z_2) = \Phi(z_2)-\Phi(z_1)$ $P(-z_1<Z< z_2) = \Phi(z1)+\Phi(z2)$ $P(Z<-z_1) = 0{,}5-\Phi(z_1)$

9.) MATEMATIČNO UPANJE (PRIČAKOVANA VREDNOST) SLUČAJNE SPREMENLJIVKE

Število E(X) je aritmetična sredina sluč. spr. in ga imenujemo matematično upanje ali pričakovana vrednost sluč. spr. X. X: $\binom{x_1\ x_2\ \cdots\ x_m}{f_1\ f_2\ \ldots\ f_m}$, kjer so f_i frekvence posameznih vrednosti. Če poskus ponovimo velikokrat, se relativne frekvence običajno ustalijo pri verjetnostih $p_i=P(X=x_i)$. Zato se pri velikem št. poskusov aritmetična sredina sluč. spr. X običajno ustali pri vrednosti $E(X) = \sum_{i=1}^m p_i x_i$
Diskretna sluč. spr. ima matematično upanje definirano kot $E(X) = \sum_{x_k}x_kP(X=x_k)$
Zvezna sluč. spr. z gostoto $p_X(x)$ ima matematično upanje definirano kot $E(X) = \int_{-\infty}^{\infty}xp(x)dx$

10.) DISPERZIJA SLUČAJNE SPREMENLJIVKE (BINOMSKA, NORMALNA)
Disperzija ali varianca sluč. spr. X meri razpršenost sluč. spr. in je definirana kot: $D(X) = E(X^2) – E(X)^2$
Za diskretno sluč. spr. z m vrednostmi je $D(X) = \sum^{m}_{i=1} (x_i - E(X))^2 p_i$
Za zvezno sluč. spr. pa $D(X) = \int^{\infty}_{-\infty}(x – E(X))^2 p(x)dx$ Kvadratni koren iz disperzije je **stand.odkl.**
Za disp velja: $D(X+b) = D(X)$ in $D(aX) = a^2D(X)$, kjer sta a, b konst.Disp je vedno >0,

11.) KOVARIANCA, PEARSONOV KOEFICIENT KORELACIJE
Kovarianca K(X,Y) sluč. spr. X in Y je določena z izrazom $K(X,Y) = E(XY) – E(X)E(Y)$. Če obstajata D(X) in D(Y),
obstaja tudi K(X,Y) in velja $|K(X,Y)| \leq \sqrt{D(X)D(Y)}$. Spr. X in Y sta nekorelirani ntk je $K(X,Y) = 0$.
Korelacijski koeficient sluč. spr. X in Y je določen z izrazom $r(X,Y) = K(X,Y) / \sigma X \sigma Y$. $R(X,Y) = 0$ ntk sta X in Y nekorelirani.

12.)SREDINE (ARITMETIČNA, GEOMETRIJSKA, HARMONIČNA)
Aritmetična sredina ali povprečje je vsota vseh vrednosti deljena s št. enot v populaciji.
Primerna je za številske, približno normalno porazdeljene spremenljivke. $\mu = \sum^{N}_{i=1}x_i / N$
Če spremenljivka doseže k različnih vrednosti s pripadajočimi frekvencami f_i, aritm. sredino
izračunamo kot $\mu = \sum^{r}_{i=1}f_ix_i / N$ in ji pravimo tehtana oz. uravnotežena aritmetična sredina, frekvencam pa uteži.
Geometrijska sredina je enaka N-temu korenu iz produkta N vrednosti št. spremenljivke pri pogoju $x_i > 0$. $G = \sqrt[N]{x_1x_2\ldots x_N}$
Harmonična sredina je enaka recipročni vrednosti aritmetične sredine, izračunane iz recipročnih vrednosti spremenljivk. $H = N / (1/x_1+1/x_2+\ldots+1/x_N)$ Velja $H\leq G\leq \mu$

13.)ČEBIŠEVA NEENAKOST Oceni nam kakšna je verjetnost, da slučajna spremenljivka veliko razlikuje od matematičnega upanja.

14.)STATISTIKA (OSNOVNI POJMI, VRSTE SPREMENLJIVK, TIPI ANALIZ) ENOTA – posamezni preučevani element populacije.
POPULACIJA – množica vseh statistično preučevanih elementov. VZOREC – podmnožica populacije, na osnovi katere ponavadi sklepamo o lastnostih celotne populacije. SPREMENLJIVKA – lastnost enot , ki nas v konkretnem raziskovanju zanimajo. Vrste spremenljivk glede na tip izražanja vrednosti:
OPISNE (atributivne) – vrednosti lahko opišemo le z besedami (npr. poklic, uspeh)
ŠTEVILSKE (numerične) – vrednosti lahko izrazimo s števili (npr. starost)
Vrste spremenljivk glede na tip merjenja: NOMINALNE – vrednosti lahko le razlikujemo med seboj (=, ≠); ostale primerjave niso smiselne; npr. barva las, spol, je_davčni_zavezanec
ORDINALNE – vrednosti lahko uredimo od najmanjše do največje; razlika med vrednostmi ni smiselna; npr. uspeh v šoli, višina ljudi INTERVALNE – smiselna je primerjava razlike med pos.vred , razmerja pa niso smiselna; npr. temperatura v C, stanje na bančnem računu RAZMERNOSTNE – smiselna je tudi primerjava razmerij med posameznima vrednostma; npr. temperatura v K, masa, starost
Tipi statističnih analiz: OPISNA statistika – statistična analiza zbranih podatkov brez težnje, da bi iz teh podatkov posploševali čez njihov obseg. INFERENČNA statistika – statistično sklepanje iz vzorca na populacijo. Ocenjevanje značilnosti populacije (intervali zaupanja), preverjanje domnev (testiranje hipotez).
Statistična analiza je: UNIVARI (analiza 1 spr.), BIVARIATNA (analiza2 spr.) in MULTIVAR (analiza več spr.)

15.)OPISNA STATISTIKA (KORAKI STATIANALIZE TER UREJANJE IN PRIKAz PODATKOV)
Koraki statistične analize: 1. določitev vsebine in namena statističnega preučevanja; opredelitev predmeta opazovanja (enote in populacije) in vsebine opazovanja (spremenljivke)
2. statistično opazovanje; vrste opazovanj: opazovanje cele populacije (npr. popisi, tekoče registracije), opazovanje vzorca (npr. ankete) 3. osnovna obdelava; urejanje, razvrščanje podatkov ter izračun osnovnih statističnih karakteristik 4. analitična obdelava Urejanje in prikazovanje podatkov:
Število vseh možnih vrednosti preučevane spr. je lahko preveliko za pregledno prikazovanje podatkov. Zato sorodne vrednosti razvrstimo v skupine. Posamezni skupini priredimo ustrezno reprezentativno vrednost spr. Skupine vrednosti morajo biti določene enolično – vsaka enota s svojo vrednostjo je lahko uvrščena v eno in samo eno skupino vrednosti.

16.) FREKVENČNA PORAZDELITEV
Frekvenčna porazdelitev spr. je tabela, ki jo določajo vrednosti ali skupine vrednosti in njihove frekvence. Če je spr. vsaj ordinalnega značaja, vrednosti uredimo od najmanjše do največje.
ŠIRINA i-tega razreda je $d_i = x_{i,min} – x_{i,max}$ (spodnja in zgornja meja i-tega razreda)
SREDINA i-tega razreda je $x_i = (x_{i,min} + x_{i,max}) / 2$
KUMULATIVA je frekvenca do spodnje meje določenega razreda. $F_{i+1} = F_i+f_i$
RELATIVNA frekvenca in kumulativa $f_i\%=(f_i /N)100$ $F_i\%=(F_i /N)100$

17.) OPISNA STATISTIKA (DEF. KVANTILA, SREDNJE VREDNOSTI, MEDIANE, MODUSA)
Ranžirna vrsta – enote z ustreznimi vrednostmi spr. uredimo od tiste z min vrednostjo do tiste z max.
Rang R – vsaki enoti v ranžirni vrsti priredimo zaporedno mesto.
Kvantilni rang P – pove, na katerem delu celotnega ranžirnega razmika leži določena enota oz. koliki del celotnega razmika ima manjše vrednosti od dane. P=(R-0,5)/N
Kvantil – je vrednost spremenljivke, ki pripada določenemu kvantilnemu rangu. Običajni kvantili so: mediana, kvartil, decil, centil. **Srednja vrednost** – reprezentativna vrednost spremenljivke, okoli katere se ponavadi gostijo enote. Čim bolj vrednosti variirajo, tem bolj se vrednosti odklanjajo od srednje vrednosti in tem slabše le-ta predstavlja spremenljivko. Vrste srednjih vrednosti: mediana M_e, modus M_o, aritmetična sredina μ, geometrijska sredina G, harmonična sredina H.
Mediana – je tista vrednost spr., od katere je enako vrednosti manjših in večjih od nje. Primerna za ordinalne spr. Dobimo jo tako, da podatke uredimo v naraščajočem vrstnem redu, nato pa vzamemo srednji element, če je podatkov liho št. oz. izračunamo povpr.rednjih elementov, če je podatkov sodo št.
Modus – je vre spr., ki se v populaciji najpogosteje pojavlja. Lahk več, lahknobenega. Primern nominalne spr.

18.) OPISNA STATISTIKA (STANDARDIZACIJA) Denimo, da vsaki vrednosti x_i spr. X odštejemo njeno
aritmetično sredino μX in delimo z njenim standardnim odklonom σX: $z_i = (x_i – \mu X) / \sigma X$ Vrednosti z $_i$ imenujemo standardizirana vrednost. Standardizirana spr. ima aritmetično sredino 0 in standardni odklon 1.

19.) MERE RAZPRŠENOSTI (POVP. ABSOLUTNI ODKLON, VARIANCA, STANDARDNI ODKLON)
Variacijski razmik: $R = x_{max} - x_{min}$ **Kvartilni odklon**: $Q = (Q_3 – Q_1) / 2$ kjer sta Q_3 in Q_1 kvartila
Povp. absolutni odklon: $AD_\mu = 1/N\sum^N_{i=1}|x_i - \mu|$ od aritmetične sredine in
$AD_{Me} = 1/N\sum^N_{i=1}|x_i - M_e|$ od mediane **Varianca oz. disperzija**: $\sigma^2 = 1/N\sum^N_{i=1}(x_i – \mu)^2$
Standardni odklon: $\sigma = \sqrt{\sigma^2}$

20.)MERE ASIMETRIJE IN SPLOŠČENOSTI, CENTRALNI MOMENTI

V primeru unimodalne porazdelitve spr., ki je bolj ali manj asimetrična in sploščena, je potrebno izračunati še stopnjo asimetrije in sploščenosti, da bi jo čim bolje opisali. **Meri asimetrije** – razlike med sredinami so tem večje, čim bolj je porazdelitev simetrična. $KA_{Mo} = (\mu - M_o) / \sigma$ in $KA_{Me} = 3(\mu - M_e) / \sigma$ (>0 asim. v desno, =0 sim., <0 asim. v levo)

Mera sploščenosti – merimo s pomočjo kvantilov (kvartilov in decilov)

$KS = 1{,}9(Q_3 - Q_1)/(D_9 - D_1)$ (<1 koničasta, =1 normalna, >1 sploščena)

Centralni moment – $m_l = 1/N\sum^{N}_{i=1}(x_i - \mu)^l$ $m_1 = 0$ $m_2 = \sigma^2$ **Koeficient asimetrije** – $g_1 = m_3 / m^{3/2}_2$

Koeficient sploščenosti – $g_2 = (m_4 / m^2_2) - 3$

21.)RELATIVNE MERE RAZPRŠENOSTI (RELATIVNI VARIACIJSKI RAZMIK, RELATIVNI KVARTILNI ODKLON, RELATIVNI POVP. ABSOLUTNI ODKLON, KOEFICIENT VARIACIJE)

Absolutne mere deljene z ustrezno srednjo vred..**Relativni var. razmik**: $(x_{max} - x_{min})2/ (x_{max} + x_{min})$

Relativni kvartilni odklon: $(Q_3 - Q_1) / 2M_e$ **Relativni povp. absolutni odklon**: AD_{Me} / M_e

Relativni standardni odklon (koeficient variacije): $KV = \sigma/\mu$

22.) PORAZDELITVE VZORČNIH STATISTIK (ARIt. SREDI, DELEŽEV, RAZLIKE ARITM. SREDIN IN DELEŽEV)

Imamo populacijo N enot, izbiramo n enot v slučajni vzorec. Pri tem ima vsaka enota enako verjetnost, da bo izbrana v vzorec t.j. 1/N. Vzorec s ponavljanjem – izbrane enote vračamo v populacijo. Vzorec brez ponavljanja – izbranih enot ne vračamo.

Populacija vseh možnih vzorcev so vsi izbrani vzorci iz populacije. Teh je v primeru ponavljanja N^n in $(^N{}_n)$ v primeru brez ponavljanja.

Porazdelitev aritmetičnih vzorčnih sredin – Denimo, da se spr. X na populaciji porazdeljuje normalno N(μ,σ). Na vsakem vzorcu (s ponavljanjem) izračunamo vzorčno aritmetično sredino X'. Porazdelitev vzorčnih aritmetičnih sredin je normalna, kjer je: matema tično upanje E(X') = μ, standardni odklon SE(X') = σ/√μ

23.)SKLEPANJE IZ VZORCA NA POPULACIJO (CENILKE: NEPRISTRANSKE IN DOSLEDNE)

Vzorčna aritmetična sredina X' je ocena populacijske aritmetične sredine μ. Vzorčno aritm. sredino imenujemo tudi **cenilka** populacijske aritmetične sredine μ. Vrednosti cenilke se od ocenjevanega parametra bolj ali manj odklanjajo. Rečemo,da je cenilka parametra dobra, če ima nekaj dobrih lastnosti: **Nepristranska cenilka** – povprečje vseh vzorčnih ocen (matematično upanje cenilke) je enako ocenjevanemu parametru. Nepristranska je ker velja E(X')= μ.

Dosledna cenilka – z večanjem vzorca se vzorčna ocena bliža parametru.

24.) **INTERVALI ZAUPANJA**

Denimo, da s slučajnim vzorcem ocenjujemo parameter γ. Poskušamo najti statistiko g, ki je nepristranska E(g) = γ in se na vseh možnih vzorcih vsaj približno normalno porazdeljuje s standardno napako SE(g). Nato poskušamo najti interval, v katerem se bo z dano gotovostjo 1-α nahajal ocenjevani parameter: $P(a < \gamma < b) = 1-\alpha$ a in b sta spodnja in zgornja meja zaupanja, α pa je verjetnost tveganja. Ta interval imenujemo **interval zaupanja** in ga interpretiramo: z verjetnostjo tveganja α se parameter γ nahaja v tem intervalu. $H(z_{\alpha/2}) = 0{,}5-\alpha/2$

25.) PREVERJANJE DOMNEV (TESTIRANJE HIPOTEZ)
Postavimo domnevo o vrednosti parametra Π – delež enot z določeno lastnostjo na populaciji. Domneva H: Π_H=… Tvorimo slučajne vzorce velikosti n in na vsakemu vzorcu določimo vzorčni delež p (delež enot z določeno lastnostjo na vzorcu). Ob predpostavki, da je domneva pravilna, vemo, da se vzorčni delež porazdeljuje približno normalno: $N(\Pi_H,\sqrt{}(\Pi_H(1-\Pi_H)/n)$ Vzemimo en vzorec z vzorčnim deležem p. Ta se lahko bolj ali manj razlikuje od Π_H. Če se zelo razlikuje, lahko podvomimo o resničnosti domneve Π_H. Zato okoli Π_H naredimo območje sprejemanja domn. in izven njega območje zavračanja domneve.
Napaka I. reda – napako storimo, če zavrnemo pravilno hipotezo.
Napaka II. **reda** – napako storimo, če sprejmemo napačno hipotezo.
Postopek preverjanja hipotez – postavimo ničelno H_0 in osnovno H_1 hipotezo o parametrih; za parameter poiščemo dobro cenilko in njeno porazdelitev ali porazdelitev ustrezne statistike; izberemo stopnjo značilnosti in na njeni osnovi določimo kritično območje; na vzorčnih podatkih izračunamo eksperimentalno vrednost statistike; naredimo sklep – če eksperimentalna vrednost pade v kritično območje ničelno hipotezo zavrnemo, če ne pade v kritično območje nimamo dovolj dokazov za zavrnitev hipoteze.

26.)OCENJEVANJE PARAMETROV Z MAJHNIMI VZORCI (STUDENTOVA t-PORAZDELITEV, X-KVADRAT)
Če se spr. X porazdeljuje na populaciji normalno in je populacijski standardni odklon poznan, potem za vsako velikost vzorca velja, da se vzorčne aritmetične sredine porazdeljujejo normalno X': $N(\mu,\sigma/\sqrt{n})$ oz. $Z = (X'-\mu)/(\sigma/\sqrt{n})$ N(0,1). Če se spr. X ne porazdeljuje na populaciji normalno in je standardni odklon poznan, potem za velike vzorce (n>30) velja, da se vzorčne aritmetične sredine porazdeljujejo približno normalno X': $N(\mu,\sigma/\sqrt{n})$ oz. $Z = (X'-\mu)/(\sigma/\sqrt{n})$ N(0,1). Če se spr. X ne porazdeljuje na populaciji normalno in standardni odklon ni poznan, potem za velike vzorce (n>30) velja, da se vzorčne aritm sred. porazdeljujejo približno normalno X': $N(\mu,s/\sqrt{n})$ oz. $Z = (X'-\mu)/(s/\sqrt{n})$ N(0,1).
Če se spr. X ne porazdeljuje na populaciji normalno in standardni odklon ni poznan, potem tudi za majhne vzorce (n<30) velja, da se statistika $t = (X'-\mu)/(s/\sqrt{n})$ t(n-1) porazdeljuje po **Studentovi t-porazdelitvi** z (n-1) rostostnimi stopnjami. Naj bo spremenljivka Z porazdeljena standardno normalno. Tedaj pravimo da je spremenljivka $\chi^2=Z^2$ porazdeljena $\boldsymbol{\chi^2}$ z eno prostostno stopnjo. $\chi^2\sim \chi^2(1)$ Naj bodo spremenljivke $Z_1,Z_2,\ldots Z_n$ porazdeljene standardno normalno. Tedaj pravimo, da je spremenljivka $\chi^2=Z^2_1+ Z^2_2+\ldots+Z^2_n$ porazdeljena $\boldsymbol{\chi^2}$ z n stopnjami prostosti. $\chi^2\sim \chi^2(n)$

27.)DOLOČANJE TRENDA Z METODO NAJMANJŠIH KVADRATOV Trend lahko obravnavamo kot posebn i primer regresijske funkcije, kjer je neodvisna spremenljivka čas T. Če je trend $X_T = f(T)$ lahko parametre trenda določimo z **metodo najmanjših kvadratov** $\sum^n_{i=1}(X_i-X_{iT})^2 = \min$
Druge metode določanja trenda: prostoročno, metoda drsečih sredin,…

28.)REGRESIJA (REGRESIJSKA PREMICA Z METODO NAJMANJŠIH KVADRATOV)
Pri regresiji skušamo ugotoviti, kakšen vpliv ima neodvisna spr. X na odvisno Y. V ta namen postavimo regresijski model Y' = f(X). V primeru linearne regresije je Y' = a+bX. Določiti moramo parametra a in b, da se bo regresijska premica kar najbolje prilegala podatkom. Izkaže se, da je $Y' = \mu_Y+C_{XY}/\sigma^2_X(X-\mu_X)$ pri čemer sta μ_X in μ_Y povprečji spremenljivk X in Y, σ_X je standardni odklon spr. X, C_{XY} pa kovarianca obeh spr. $C_{XY} = 1/N\sum^N_{i=1}((x_i-\mu_X)(y- \mu_Y))$Ocena parametrov reg funkc z metodo najm.kv $\sum^N_{i=1}e^2_i = \sum^N_{i=1}(y_i-y'_i)^2 = \min$