1. **Srednje vrijednosti.**

Srednja vrijednost je konstanta koja ima za cilj na reprezentativan način predstaviti niz varijabilnih podataka numeričkoga niza. Vrste:

1.Aritmetička sredina,
Ona predstavlja jednaki dio vrijednosti numeričkog obilježja koji otpada na jednu jedinicu statističkog skupa.

$$\overline{X} = \frac{\sum_{i=1}^{N} x_i}{N} \quad ; \ i = 1,2,3,...,N$$

Aritmetička sredina nalazi se između najmanje i najveće vrijednosti obilježja.
Zbroj odstupanja pojedinih vrijednosti obilježja od aritmetičke sredine jednak je nuli: $\sum_{i=1}^{N}\left(x_i - \overline{x}\right) = 0$

Zbroj kvadrata odstupanja vrijednosti obilježja pojedinih elemenata statističkog skupa ima minimalnu vrijednost: $\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2 \leq \sum_{i=1}^{N}\left(x_i - a\right)^2$

Vagana (ponderirana) aritmetička sredina

$$\overline{X} = \frac{\sum_{i=1}^{k} f_i \cdot x_i}{\sum_{i=1}^{k} f_i}$$

2.Geometrijska sredina
Geometrijska sredina definira se ovako za negrupirane nizove:

$$G = \sqrt[N]{\prod_{i=1}^{N} X_i}$$

Geometrijska sredina uvijek je manja od aritmetičke sredine.

3.Harmonijska sredina,
Harmonijska sredina predstavlja recipročnu vrijednost aritmetičke sredine recipročnih vrijednosti iz kojih se ona izračunava.

$$H = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \frac{f_i}{x_i}}$$

4.Medijan,
Medijan je *pozicijska srednja vrijednost* koja statistički niz dijeli na dva jednaka dijela. Podatke treba poredati od najmanjega prema najvećemu ili obrnuto. Medijan se računa za numeričke i redosljedne nizove.

5.Mod.
Mod ili dominanta predstavlja najčešću vrijednost obilježja. Može se računati za sve vrste nizova. Prednost medijana u odnosu na aritmetičku sredinu je u tome što ne reagira na ekstremne vrijednost.

2. Mjere disperzije.

Pod pojmom disperzije podrazumijevamo raspršenost vrijednosti numeričkoga obilježja. Mjere disperzije služe za ocjenjivanje reprezentativnosti srednje vrijednosti obilježja.

Najčešće u uporabi su ove mjere disperzije:

Apsolutne mjere disperzije: raspon varijacije obilježja, prosječno apsolutno odstupanje, varijanca i standardna devijacija,
interkvartil.
Relativne mjere disperzije: koeficijent varijacije, koeficijent kvartilne devijacije.

Prosječno apsolutno odstupanje (“**M**ean **A**bsolute **D**eviation”) dobiva se kao aritmetička sredina apsolutnih vrijednosti odstupanja od aritmetičke sredine vrijednosti obilježja:

$$MAD = \frac{\sum_{i=1}^{N} \left| X_i - \overline{X} \right|}{N}$$

Varijanca je srednje kvadratno odstupanje numeričkih vrijednosti obilježja od aritmetičke sredine. **Standardna devijacija** je pozitivni korijen iz varijance i predstavlja apsolutnu mjeru disperzije u prvome stupnju.
Varijanca za negrupirane podatke računa se preko sljedećega izraza:

$$\sigma^2 = \frac{\sum_{i=1}^{N} \left(X_i - \overline{X} \right)^2}{N} \quad \text{ili} \quad \sigma^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\overline{x} \right)^2$$

Varijanca za grupirane nizove:

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i \cdot \left(x_i - \overline{x} \right)^2}{\sum_{i=1}^{k} f_i} \quad \text{ili} \quad \sigma^2 = \frac{\sum_{i=1}^{k} f_i \cdot x_i^2}{\sum_{i=1}^{k} f_i} - \left(\overline{x}^2 \right)$$

Standardna devijacija je pozitivni korijen iz varijance: $\sigma = +\sqrt{\sigma^2}$
Interkvartil predstavlja apsolutnu mjeru disperzije srednjih 50% jedinica u statističkome skupu. Time interkvartil izbjegava ekstremne vrijednosti obilježja i isključuje 25% najmanjih i 25% najvećih vrijednosti.

$$I_q = Q_3 - Q_1$$

Kvartili zajedno s medijanom dijele statistički skup na četiri jednaka dijela.
Koeficijent varijacije relativna je mjera disperzije:

$$V = \frac{\sigma}{\overline{x}} \cdot (100)$$

Koeficijent kvartilne devijacije relativna je mjera disperzije srednjih 50% jedinica u statističkome nizu.

$$V_q = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$ Njegova vrijednost kreće se između 0 i 1.

U ZADACIMA još RADILI:

Pearsonov koeficijent asimetrije dobiva se kao omjer trećega centralnoga momenta i standardne devijacije prethodno dignute na treću potenciju.

$$\alpha_3 = \frac{\mu_3}{\sigma^3} \qquad \mu_3 = \frac{\sum_{i=1}^{k} f_i \cdot (X_i - \overline{X})^3}{\sum_{i=1}^{k} f_i}$$ kod negrupiranih: $$\mu_3 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^3}{N}$$

Kod simetričnih distribucija vrijednost Pearsonovoga koeficijenta asimetrije jednaka je nuli.Kod desnostrane (pozitivne) asimetrije veći je od nule, dok je kod ljevostrane (negativne) asimetrije manji od nule.
U većini slučajeva: $-2 \le \alpha_3 \le +2$
Pearsonova mjera asimetrije definira se na temelju odnosa aritmetičke sredine i moda na sljedeći način:

$$S_k = \frac{\overline{X} - M_o}{\sigma}$$

Kao jedna od karakteristika numeričkih vrijednosti obilježja distribucije može se uvesti još i zaobljenost vrha (kurtoza). U tu svrhu se upotrebljava koeficijent zaobljenosti vrha distribucije sa sljedećim izrazom:

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \qquad \mu_4 = \frac{\sum_{i=1}^{k} f_i \cdot (x_i - \overline{x})^4}{\sum_{i=1}^{k} f_i}$$

3. Pojam vjerojatnosti. Adicijski i multiplikacijski teorem. Bernoullijev zakon velikih brojeva.

Teorija vjerojatnosti čini osnovu statističkoga zaključivanja. Vjerojatnost je spoj filozofijskoga poimanja slučaja i matematike.
Prema klasičnoj definiciji vjerojatnost realizacije slučajnoga događaja A jednaka je omjeru broja (za njega!) povoljnih ishoda i svih mogućih ishoda:

$$P(A) = \frac{m(A)}{n}$$

vjerojatnost je mjera slučaja koja se kreće između 0 i 1.
Vjerojatnost da se slučajni događaj (A) **ne realizira** jednaka je omjeru broja za njega nepovoljnih ishoda i broja svih mogućih ishoda:

$$p(\overline{A}) = \frac{n - m(A)}{n}$$ stoga je: $P(A) + P(\overline{A}) = 1$

Kod vjerojatnosti “a priori” **unaprijed** je poznat broj povoljnih i broj svih mogućih ishoda.
Ukoliko vjerojatnost realizacije slučajnoga događaja A nije poznata unaprijed, može se izračunati tzv. "vjerojatnost a posteriori".

$$P(A) = p\lim_{n\to\infty}\frac{m(A)}{n}$$

Izraz p lim se čita "granična vrijednost po vjerojatnosti" da se razlikuje od limesa u linearnoj algebri. Gornji izraz predstavlja **Bernoullijev zakon velikih brojeva**. Vjerojatnost "a posteriori" jednaka je graničnoj vrijednosti relativne frekvencije kada broj pokusa teži u beskonačnost.

Ako dva slučajna događaja ne mogu nastupiti istodobno, kažemo da se ti događaji međusobno isključuju. Vjerojatnost realizacije jednoga ili drugoga događaja jednaka je zbroju vjerojatnosti realizacije jednoga i vjerojatnosti realizacije drugoga događaja. $P(A \text{ ili } B) = P(A) + P(B)$

Ukoliko se slučajni događaji A i B ne isključuju, tada se vjerojatnost realizacije slučajnoga događaja A ili događaja B dobiva na sljedeći način:

$$P(A \text{ ili } B) = P(A) + P(B) - P(AB)$$

P(AB) predstavlja vjerojatnost da slučajni događaji A i B nastupe istovremeno.Teorem se može poopćiti na više događaja.

Ukoliko se događaji A i B međusobno ne isključuju, i ako vjerojatnost realizacije jednoga događaja ne zavisi o vjerojatnosti realizacije drugoga događaja, tada je vjerojatnost istovremene realizacije događaja A i događaja B jednaka umnošku vjerojatnosti realizacije događaja A i događaja B: $P(A \text{ i } B) = P(A) \cdot P(B)$
Za takove događaje kažemo da su (stohastički) nezavisni.

4. Uvjetna vjerojatnost. Bayesov teorem.
Ukoliko je realizacija događaja A uvjetovana prethodnom realizacijom događaja B, radi se o uvjetnoj ili KONDICIONALNOJ vjerojatnosti događaja A.

$$P(A/B) = \frac{P(AB)}{P(B)} \quad , \; P(B) > 0 \quad \text{odnosno} \quad P(B/A) = \frac{P(AB)}{P(A)} \quad , \; P(A) > 0$$

Kod nezavisnih događaja vrijedi sljedeće: $P(A/B) = P(A)$
Bayesov teorem:
Ako se događaj B realizira samo tada kada nastupi jedan od n disjunktnih događaja A1, A2, ..., An za koje je $\sum_{i=1}^{n} P(A_i) = 1$ onda se vjerojatnost događaja B dobiva po formuli potpune vjerojatnosti: $P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B/A_i)$

$$P(A_k/B) = \frac{P(A_k) \cdot P(B/A_k)}{\sum_{i=1}^{n} P(A_i) \cdot P(B/A_i)} \qquad \sum_{i=1}^{n} P(\mathrm{A_i}) = 1$$

5. Diskontinuirana slučajna varijabla. Svojstva i teorijske distribucije diskontinuirane slučajne varijable.

Diskontinuirana (ili diskretna) slučajna varijabla je takova varijabla koja može poprimiti najviše prebrojivo beskonačno mnogo vrijednosti s određenom vjerojatnošću: $P(x_1), P(x_2), ...$
pri čemu mora biti:

$$\sum_{i=1}^{\infty} P(x_i) = 1 \quad \text{i} \quad P(x_i) \geq 0 \quad \forall i$$

Uređeni skup parova: $\{x_i, P(x_i)\},\ i = 1,2,3,...$
naziva se distribucija slučajne varijable X. Zakon po kojem svakoj vrijednosti slučajne varijable X pripada vjerojatnost P(xi) naziva se ***zakon vjerojatnosti*** slučajne varijable X.

Funkcija distribucije slučajne varijable X predstavlja vjerojatnost da slučajna varijabla X ne premaši neku određenu vrijednost:

$$F(x_k) = \sum_{i=1}^{k} P(x_i) = P(X \leq x_k)$$

Očekivanje diskontinuirane slučajne varijable jednako je zbroju umnožaka vrijednosti varijable X i odgovarajućih vjerojatnosti:

$$E(X) = \sum_{i=1}^{\infty} x_i \cdot p(x_i) = \mu$$

Varijanca diskontinuirane slučajne varijable X je očekivanje kvadratnoga odstupanja vrijednosti varijable X od njenoga očekivanja:
$V(X) = E(X - \mu)^2$ odnosno $V(X) = E(X^2) - \left[E(X)\right]^2$

- Teorijske distribucije diskontinuirane slučajne varijable
Binomna distribucija
Ako je vjerojatnost da nastupi neki slučajni događaj **poznata i uvijek ista tijekom izvođenja pokusa** može se izračunati vjerojatnost da se slučajna varijabla *X* realizira *x* puta u *n* pokusa. U tome slučaju kažemo da se diskontinuirana slučajna varijabla X ravna prema binomnoj distribuciji. Binomna distribucija ima sljedeći zakon vjerojatnosti:

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x} \quad x = 0,1,2,3,...,n$$

n = broj pokusa,
x = broj povoljnih ishoda u n pokusa,
p = vjerojatnost realizacije slučajnoga događaja,
n - x = broj nepovoljnih ishoda u n pokusa.

$$(E(x) = n \cdot p, \quad \sigma^2 = n \cdot p \cdot q, \quad \alpha_3 = \frac{q - p}{\sqrt{n \cdot p \cdot q}} \quad \alpha_4 = 3 + \frac{1 - 6 \cdot p \cdot q}{n \cdot p \cdot q})$$

Poissonova Distribucija
Ako je vjerojatnost slučajnoga događaja veoma malena i konstantna tijekom izvođenja pokusa, umjesto binomne distribucije može se koristiti Poissonova

distribucija. U tome slučaju broj pokusa raste u beskonačnost, ali očekivana vrijednost ostaje konstantna. Što je vrijednost od n veća, a vrijednost od p manja, to je aproksimacija bolja. Izraz za Poissonovu distribuciju je sljedeći:

$$P(X = x) = \frac{(np)^x \cdot e^{-np}}{x!} \qquad x = 0,1,2,3,...\infty$$

odnosno, s obzirom da je funkcija poprima oblik:

$$p(X = x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \qquad x = 0,1,2,3,...,\infty$$

$$E(X) = \mu \quad \mathrm{V(X)} = \mu \quad \alpha_3 = \frac{1}{\sqrt{\mu}} \quad \alpha_4 = 3 + \frac{1}{\mu}$$

Hipergeometrijska Distribucija

Osnovni skup sastoji se iz dva dijela: skup elemenata koji imaju neko obilježje A i skup elemenata koji to obilježje nemaju. U uzorak se bira određeni broj elemenata osnovnoga skupa. Slučajna varijabla je broj jedinica u uzorku koje imaju određeno obilježje A. Ukoliko se jedinice osnovnoga skupa, koje su jednom uzete u uzorak ne vraćaju ponovno u osnovni skup, radi se o hipergeometrijskoj distribuciji:

$$P(X = x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{\mathrm{N}}{\mathrm{n}}} \quad , \quad x = 0,1,2,...,n$$

P(X) = zakon vjerojatnosti slučajne varijable X,
M = broj elemenata u osnovnom skupu s obilježjem A,
N - M = broj elemenata u osnovnom skupu koji nemaju obilježje A,
n - x = broj elemenata u uzorku koji nemaju obilježje A,
N = broj elemenata u osnovnom skupu,
n = broj elemenata u uzorku,
x = broj elemenata u uzorku koji imaju obilježje A.

$$E(X) = n \cdot \frac{M}{N} \qquad V(X) = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$$

Jednolika Distribucija

Ako slučajna varijabla X poprima s istom vjerojatnošću bilo koju od n vrijednosti, kažemo da X ima jednoliku distribuciju:

$$P(X = x_i) = \frac{1}{n}, \qquad \forall x_i \in R_x$$

6. Dvodimenzionalna distribucija vjerojatnosti. Marginalna distribucija vjerojatnosti.

Diskontinuirana slučajna varijabla X može poprimiti vrijednosti $x_1, x_2, x_3, ..., x_k$, dok diskontinuirana slučajna varijabla Y može istovremeno poprimiti vrijednosti $y_1, y_2, y_3, ..., y_m$.

Vjerojatnost da slučajna varijabla X poprimi vrijednost xi, a istovremeno slučajna varijabla Y poprimi vrijednost yj označava se ovako:

$$P(X = x_i, Y = y_j) = P(x_i, y_j)$$

Budući da se radi o distribuciji vjerojatnosti, moraju biti zadovoljeni sljedeći uvjeti:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j) = 1 \qquad P(x_i y_j) \geq 0$$

*****za zadatke triba znat (primjer):

$$E(X/Y=1) = \sum_i x_i \cdot P(x_i / y=1)$$

$$V(X/Y=1) = \sum_i x_i^2 \cdot P(x_i / y=1) - \left[E(X/Y=1)\right]^2$$

Skup svih uređenih parova $\{(x_i, y_j); P(x_i, y_j)\}$ sačinjava dvodimenzionalnu distribuciju slučajne varijable (X,Y).

Marginalne Distribucije Diskontinuirane Slučajne Varijable

Kod marginalnih distribucija primjenjuje se adicijski teorem. Traži se vjerojatnost da slučajna varijabla X poprimi neku vrijednost Xi bez obzira na to koju će vrijednost poprimiti slučajna varijabla Yj. Slično tomu, marginalna distribucija slučajne varijable Y predstavlja vjerojatnosti da varijabla Y poprimi vrijednost Yj bez obzira koju vrijednost poprima slučajna varijabla X.

$$P(x_i) = \sum_{j=1}^{m} P(x_i, y_j) \qquad P(y_j) = \sum_{i=1}^{n} P(x_i, y_j)$$

Vjerojatnost da slučajna varijabla X poprimi neku vrijednost xi **pod uvjetom** da je varijabla Y poprimila neku vrijednost yj naziva se uvjetnom ili kondicionalnom distribucijom vjerojatnosti:

$$P(x_i / y_j) = \frac{P(x_i, y_j)}{\sum_{j=1}^{m} P(x_i, y_j)} = \frac{P(x_i, y_j)}{P(y_j)}$$

7. Kontinuirana slučajna varijabla. Svojstva i teorijske distribucije kontinuirane slučajne varijable.

Kontinuirana slučajna varijabla X može poprimiti neprebrojivo beskonačno mnogo vrijednosti. Kod kontinuiranih slučajnih varijabli ne računa se vjerojatnost u određenoj točki, nego nad određenim intervalom vrijednosti slučajne varijable X. Vjerojatnost da će neka kontinuirana slučajna varijabla X poprimiti neku određenu vrijednost x jednaka je nuli, ali to ne znači i nemogući događaj.

Kod kontinuirane slučajne varijable umjesto P(X) upotrebljavamo f(X). Funkcija vjerojatnosti f(X) nije vjerojatnost, ali je tom funkcijom određena vjerojatnost što pripada svakom intervalu (x_1, x_2)

Funkcija vjerojatnosti (naziva se i funkcija gustoće vjerojatnosti) kontinuirane slučajne varijable X ima sljedeća svojstva:

$$f(x) \geq 0 \qquad \forall x$$

Površina ispod krivulje vjerojatnosti jednaka je 1:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \qquad \int_{x_1}^{x_2} f(x)dx = P\{x_1 < X \leq x_2\} \qquad x_2 > x_1$$

Funkcija distribucije kontinuirane slučajne varijable X je sljedeća:

$$F(x) = \int_{-\infty}^{x} f(x)dx \qquad \frac{dF(x)}{dx} = f(x)$$

F(x) predstavlja vjerojatnost da slučajna varijabla X ne premaši neku unaprijed zadanu vrijednost x. U empirijskoj statistici ta funkcija predstavlja kumulativni niz "manje od".

Očekivanje i varijanca kontinuirane slučajne varijable:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \mu \qquad V(x) = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)dx = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - \mu^2$$

TEORIJSKE DISTRIBUCIJE

- Normalna distribucija je najvažnija distribucija u statističkoj teoriji. Funkcija gustoće vjerojatnosti normalne distribucije je sljedeća:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad x \in <-\infty, +\infty>$$

Parametri normalne distribucije su njeno očekivanje μ i varijanca σ^2.

$$Z = \frac{x-\mu}{\sigma}$$

Ovdje se radi o normalnoj distribuciji *N(0;1)* koju nazivamo standardiziranom ili jediničnom normalnom distribucijom s očekivanjem koje je jednako nuli, a varijancom jednakom jedinici.

-Studentova distribucija

Studentova distribucija česta je u primjenama u procjeni parametara i u testiranju hipoteza na osnovu uzorka. Njezin oblik ovisi o broju stupnjeva slobode ν. Područje vrijednosti varijable ***t*** je interval $(-\infty\ ;\ +\infty)$. Studentova distribucija je simetrična s obzirom na t = 0. Spljoštenija je od normalne distribucije. Ukoliko $\nu \to \infty$ Studentova distribucija teži jediničnoj normalnoj distribuciji.

Ako je Z varijabla jedinične normalne distribucije *N(0;1),* a varijabla gama distribucije s brojem stupnjeva slobode ν, onda je:

$t = \dfrac{z}{\sqrt{\dfrac{\chi^2}{\nu}}}$ varijabla Studentove distribucije s brojem stupnjeva slobode ν.

-HI-kvadrat distribucija

Ako su X1, X2, X3, ..., Xn nezavisne normalne varijable koje imaju jednaka očekivanja, E(X1) = ... = E(Xn) = μ i jednake varijance V(X1) =.....= V(Xn) =   ,
tada je:

$\chi^2 = \sum_{i=1}^{n} \left(\dfrac{x_i - \mu}{\sigma} \right)^2$ gama varijabla sa stupnjem slobode ν = n.

-F-distribucija
F- distribucija je određena s dva parametra $\nu1$ i $\nu2$ koji predstavljaju stupnjeve slobode. Ako su χ_1^2 i χ_2^2 nezavisne $\chi2$ razdiobe sa stupnjevima slobode $\nu1$ i $\nu2$ tada varijabla:

$$F = \frac{\chi_1^2 / \nu_1}{\chi_2^2 / \nu_2}$$ pripada F-distribuciji sa stupnjevima slobode $\nu1$ i $\nu2$.

8. Metoda uzoraka. Pristranost procjene na osnovu uzoraka. Nabrojite sve vrste grješaka koje se javljaju kod rada s uzorcima.

Uzorak predstavlja podskup osnovnoga statističkog skupa koji se uzima u svrhu ispitivanja obilježja elemenata osnovnoga skupa. Da bi tu svrhu uzorak mogao ispuniti, potrebno je da bude reprezentativan i da je izbor jedinica izvršen na slučajan način. Da bi se postigla slučajnost izbora jedinica u uzorak, najčešće se upotrebljavaju tablice slučajnih brojeva.
Frakcija odabiranja predstavlja omjer broja jedinica u uzorku i broja jedinica u osnovnome skupu:

$$f = \frac{n}{N}$$ *n* = broj jedinica u uzorku, *N* = broj jedinica u osnovnome skupu

9. Procjena aritmetičke sredine na osnovu uzorka.

Procjena aritmetičke sredine na osnovu uzorka naziva se točkastom procjenom aritmetičke sredine osnovnoga skupa. Da bi se dobila intervalna procjena aritmetičke sredine, potrebno je u račun uzeti varijabilitet jedinica u osnovnome skupu (izražen preko varijance osnovnoga skupa ili njene ocjene pomoću uzorka) i veličinu uzorka.
Intervalna procjena aritmetičke sredine dobiva se ovako:

$$\Pr\left\{ \hat{\bar{X}} - z \cdot Se(\bar{x}) < \bar{X} < \hat{\bar{X}} + z \cdot Se(\bar{x}) \right\} = 1 - \alpha$$

1-α naziva se intervalom povjerenja.

Standardna greška:

$$Se(\hat{\bar{X}}) = \frac{\hat{\sigma}}{\sqrt{n}}$$

odnosno za konačne skupove:

$$Se(\hat{\bar{X}}) = \frac{\hat{\sigma}}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$ Za male uzorke (n < 30) koristimo “t” iz Studentove

Određivanje veličine uzorka:

$$n' = \left[\frac{z \cdot \sigma}{greska} \right]^2$$

Korekcija za konačne osnovne skupove gdje je frakcija manja od 0.05

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

10. Procjena proporcije (relativne frekvencije) na osnovu uzorka.

$$\Pr\left\{ \hat{p} - z \cdot Se(p) < P < \hat{p} + z \cdot Se(p) \right\} = 1 - \alpha$$

Procjena proporcije na osnovu uzorka nepristrana je.
Pri tomu se standardna grješka procjene proporcije dobiva ovako:

$$Se(p) = \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$ ako je n > 30 jedinica

$$Se(p) = \sqrt{\frac{\hat{p} \cdot \hat{q}}{n-1}}$$ ako je n < 30 jedinica

p = m / n;   q=(1-p)
Korekcija za konačne skupove je ista kao i kod procjene aritmetičke sredine osnovnoga skupa.

11. Procjena varijance na osnovu uzorka.

Procjena varijance na osnovu uzorka je pristrana. Sampling distribucija varijanci ima oblik Hi-kvadrat distribucije. Nepristrana ocjena varijance osnovnoga skupa može se dobiti ovako:

$s^2 = \hat{\sigma} \cdot \frac{n}{n-1}$ s^2 predstavlja nepristranu procjenu varijance osnovnoga skupa

Intervalna procjena varijance osnovnoga skupa:

$$\Pr\left\{ \frac{n \cdot s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{n \cdot s^2}{\chi^2_{1-\alpha/2}} \right\} = 1 - \alpha$$

U koliko veličina uzorka prelazi 30 jedinica Hi-kvadrat distribucija distribucija može se aproksimirati pomoću normalne distribucije. Tada se vrijednosti Hi-kvadrata ogu dobiti pomoću slijedeće formule:

$$\chi^2_{\alpha/2} = \frac{1}{2}\left(z_{\frac{1-\alpha}{2}} + \sqrt{2 \cdot v - 1} \right)^2 \quad \chi^2_{1-\alpha/2} = \frac{1}{2}\left(-z_{\frac{1-\alpha}{2}} + \sqrt{2 \cdot v - 1} \right)^2$$