

Vjerojatnost i statistika

Božidar Ivanković

Izvanredna nastava, proljeće 2015

Ukratko

Kolegij "Vjerojatnost i statistika" sastoji se od tri povezana područja:

- 1 Kombinatorika i Vjerojatnost
- 2 Slučajne varijable
- 3 Statistika

Literatura: E. Kovač Striko, T. Fratrović, B. Ivanković: *Vjerojatnost i statistika s primjerima iz tehnologije prometa*

Rješavanje ispita

- 1 Potpis:
 - 1 Domaće zadaće
 - 2 Aktivna prisutnost
- 2 Prolaz:
 - 1 Izvanredni pismeno-usmeni kolokvij za aktivno prisutne
 - 2 Na redovnom roku najmanje dva boda iz pismenog.
- 3 Usmeni dio obavezan nakon prolaska pismenog

Algebra skupova. Vennovi dijagrami

Skupovi ($A, B, C \dots$) su zatvorene cjeline - Vennovi dijagrami. Univerzalni skup (Ω) čine svi mogući elementi. Prazan skup jedini nema elemenata.

Definicija

Operacije sa skupovima

- unija $A \cup B$,
- presjek $A \cap B$,
- komplement u univerzalnom skupu A^c ,
- razlika skupova $A \setminus B$

Zadatak

*Vennovim dijagramima pokažite identitete: $A \cap B^c = A \setminus B$,
 $(A \cap B)^c = A^c \cup B^c$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.*

Veličina skupa

Veličina skupa može biti

- prirodan broj - kardinalitet - $|A|$
- neograničena ali prebrojiva
- neograničena ali neprebrojiva
- ograničena ali neprebrojiva

Zadatak

Odredite veličine skupova

- 1 *Moguće vrijednosti ukupne potrošnje vode sutra u Zagrebu.*
- 2 *Karte iz špila koje se mogu izvući dok se ne izvuče herc bez vraćanja izvučene karte.*
- 3 *Karte iz špila koje je moguće izvući dok se ne izvuče herc, ali sa vraćanjem svake izvučene karte.*
- 4 *Temperature na kojoj će voda u loncu zakipiti.*

Formula uključivanja i isključivanja

Teorem

Za skupove A i B vrijedi $|A \cup B| = |A| + |B| - |A \cap B|$.

Primjer

U studentskoj grupi su 24 studenta. Engleski govori 20 studenata, a njemački 10. Kako je to moguće?

Zadatak

Od svih studenata upisanih na prvoj godini 40% ih je položilo Matematiku, 50% Osnove elektrotehnike i 80% Računarstvo. Nadalje, 30% studenata položilo je Matematiku i Osnove elektrotehnike, 40% Osnove elektrotehnike i Računarstvo, dok ih je 25% položilo Matematiku i Računarstvo. Ako se zna da je 20% studenata položilo sva tri predmeta, koliki postotak studenata nije položio niti jedan od navedena tri kolegija?

rješenja: bar 6 zna oba jezika, 5% niti jedan

Kombinatorika

Teorem (o uzastopnom prebrojavanju)

Neka se u uređenoj m -torki (a_1, a_2, \dots, a_m)

- prva komponenta a_1 može se birati na k_1 različitih načina,
- nakon već odabrane komponente a_1 , druga komponentu a_2 može se izabrati na k_2 različitih načina,

⋮

- sve dok se nakon svih prethodnih, posljednja može birati na k_m različitih načina,

Tada skup T svih uređenih m -torki (a_1, a_2, \dots, a_m) ima

$$k_1 \cdot k_2 \cdots k_m$$

elemenata.

Primjer i zadatak

Primjer

Za ručak su ponuđena dva glavna jela i tri deserta. Na koliko se različitih načina može objedovati tako da se uzme točno jedno glavno jelo i točno jedan desert? ($2 \cdot 3$)

Zadatak

Profesor ima tri odijela, osam košulja, 10 kravata i dva para cipela. Na koliko se različitih načina može obući? ($3 \cdot 8 \cdot 11 \cdot 2$, ne mora kravatu.)

Zadatak

U zatvoru za ručak su ponuđene dvije vrste juhe, tri vrste glavnog jela i četiri vrste kolača. Na koliko se načina može objedovati, ako korisnik KPD-a može uzeti najviše po jednu juhu, glavno jelo i kolač? ($3 \cdot 4 \cdot 5 - 1$, ako štrajka glađu, ne ruča.)

Permutacije bez ponavljanja

Zadatak

Na koliko načina mogu uvijek ista prva petorica socijalnih slučajeva stajati u redu za pučku kuhinju?

Definicija

Permutacija n -članog skupa S je svaki niz od n međusobno različitih elemenata iz S .

Teorem

Broj permutacija skupa od n elemenata iznosi $n!$.

Faktorijela je unarna operacija:

$$n! = 1 \cdot 2 \cdot \dots \cdot n.$$

Broj permutacija bez ponavljanja

Zadatak

Koliko se 4-slovnih zaporki može složiti od riječi FRKA?

Zadatak

Koliko ima 7-znamenkastih telefonskih brojeva u kojima se po jednom javljaju znamenke $\{3, 4, 5, 6, 7, 8, 9\}$?

Zadatak

Na koliko načina mogu Andrija, Blaž, Cvetko i Dragec igrati Belu u četvero? (Andrija sjedne, a ostali popunjavaju slobodna mjesta: 6)

Zadatak

Na koliko načina može 11 vitezova sjesti za okrugli stol? $10!$

Permutacije s ponavljanjem

Primjer

Koliko se zaporki može napraviti od riječi TATA?

Primjer

Koliko se PIN-ova može napraviti od PIN-a 3272?

Primjer

Na koliko se načina mogu posložiti slova u riječi KRAPINA? A u riječi OTORINOLARINGOLOGIJA?

Teorem

Permutacija s ponavljanjem ima onoliko puta manje od permutacija bez ponavljanja, koliko iznosi umnožak faktorijela elemenata koji se ponavljaju.

Varijacije bez ponavljanja

Primjer

Na koliko načina može 10 umirovljenika zauzeti 5 mjesta u tramvaju?

Definicija

Varijacija r -tog razreda skupa S od n elemenata je svaka uređena r -torka medjusobno različitih elemenata iz skupa S . Broj varijacija:

$V_n^r = \frac{n!}{(n-r)!} = n \cdot (n-1) \cdots (n-r+1)$. Džepna računaljka ima program nPr koji računa varijacije bez ponavljanja.

Varijacije s ponavljanjem

Zadatak

Na koliko se načina, teoretski, može popuniti listić sportske prognoze ako se za svaki od 13 parova smije ponuditi bilo koji od tri znaka: X, 0 ili 1?

Zadatak

Tri prijatelja kupuju Peugeote koji su ponuđeni u četiri boje. Na koliko načina oni mogu izabrati boje?

Zadatak

U grupi je 6 studenata i 4 studentice. Na koliko se načina mogu studentima svidati studentice? Na koliko se načina mogu studenti svidati studenticama?

Broj varijacija r -tog razreda skupa od n elemenata s ponavljanjem jednak je n^r

Kombinacije bez ponavljanja

Primjer

Košarkaški trener prijavio je 11 igrača. Koliko različitih petorki može biti na parketu?

Jedan način odabira petorke. Svi igrači poredaju se u vrstu. Na koliko je načina to moguće? Prvih pet se u dresovima pošalje u igru. Ostalih 6 u trenirkama sjede.

$$\frac{11!}{5!6!}$$

Isto bi bilo da trener bira kojih 6 će sjediti na klupi.

Broj kombinacija bez ponavljanja

Definicija

Kombinacija od k -tog razreda skupa od n elemenata svaki je izbor podskupa od k elemenata izuzet iz skupa od n elemenata.

Propozicija

Broj kombinacija jednak je $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Binomni koeficijent $\binom{n}{k}$ moguće je na boljim džepnim računaljkama dobiti programom nCr .

Primjer

Koliko kombinacija ima Lotto 7/39? A koliko Lotto 6/45?

Kombinacije s ponavljanjem

Primjer

Na koliko načina mogu cura i dečko pojesti tri kremšnite?

Svaki od načina je objekt koji se u teoriji skupova naziva multiskup:

- $\{c, c, c\}$ cura je pojela sve tri kremšnite
- $\{c, c, d\} = \{c, d, c\} = \{d, c, c\}$ cura dvije, dečko jednu
- $\{c, d, d\} = \dots$ dvije je pojeo dečko
- $\{d, d, d\}$

Ukupan broj je

$$\binom{2 + 3 - 1}{3}.$$

Kombinacije s ponavljanjem

Problem

Kako može pet dječaka podijeliti 12 bombona?

Definicija

Kombinacija r -tog razreda skupa od n elemenata s ponavljanjem je multiskup od r elemenata u kojem je najviše n vrsta elemenata.

Teorem

Broj kombinacija računa se po formuli: $\overline{K}_n^r = \binom{n+r-1}{r}$

Zadaci

- 1 Dobavljač naručuje pet glačala koja mogu biti ispravna ili neispravna. Na koliko načina se mogu dobiti?
- 2 Na koliko se načina mogu 3 jednake knjige razdijeliti među 12 učenika, tako da svaki učenik dobije najviše jednu knjigu?
- 3 Registarska pločica zagrebačkog registarskog područja ima u izboru 3 i 4-znamenkasti broj, te jedno ili dva slova abecede koja ne mogu biti palatali, koliko se različitih tablica može izdati?
- 4 Na nekom je šahovskom turniru odigrano 78 partija. Turnir je igran po principu da je svaki igrač odigrao sa svakim samo jedan meč. Koliko je bilo šahista na turniru?
- 5 U skupu od 100 studenata engleski jezik zna 28 studenata, njemački 30, francuski 42, engleski i njemački zna njih 8, engleski i francuski 10, njemački i francuski pet, dok sva tri jezika znaju samo tri studenta. Koliko studenata ne zna nijedan jezik?

Rješenja

- 1 *Rješenje:* 6
- 2 *Rješenje:* 1320 načina
- 3 *Rješenje:* 5 566 000
- 4 *Rješenje:* 13 učesnika
- 5 *Rješenje:* 20 studenata

Intuitivna vjerojatnost

Slučajan pokus je pokus s vrlinama

- nepoznat ishod
- poznati svi ishodi
- ponavljanje po volji

Odredite postotak šansi pojave događaja prije pokusa.

- 1 Kolika je vjerojatnost da će kod bacanja simetričnog novčića pasti "pismo"?
- 2 Kolika je vjerojatnost da će kod bacanja kocke pasti broj 6?
- 3 Kolika je vjerojatnost da će kod izvlačenja jedne karte iz špila od 32 biti izvučena karta boje "karo"?
- 4 Kolika je vjerojatnost da razlika brojeva na bačene dvije kocke bude manja od tri?

Algebra događaja

Elementarni događaj je jedan ishod slučajnog pokusa: ω

Prostor elementarnih događaja je univerzalni skup svih ishoda Ω

Događaj je skup

Nemogućim događajem naziva se

- prazan skup $\emptyset \subset \Omega$
- bilo koji skup C : $C \cap \Omega = \emptyset$

Sigurnim događajem naziva se skup Ω .

Operacije s događajima su logičke:

ili (unija); *i* (presjek); *ne* (komplement).

Intuitivna definicija vjerojatnosti

Vjerojatnost događaja - $p(A) = \frac{|A|}{|\Omega|}$.

Zadatak

Koliko je vjerojatna sedmica na lotu ako igrač odigra sistem od 23 broja?

Zadatak

Koliko je vjerojatno da mali Ivica pogodi prozor loptom ako noću napucava zid 10×5 metara na kojem je prozor 1.20×1 metar.

Zadatak

Policiji je dojavljeno da noću od dva do tri manijak prolazi kroz crveno. Koliko je vjerojatno da policija ulovi izgređnika ako

- *čeka u zasjedi od pola tri do petnaest do tri.*
- *u dva i petnaest prođu raskršćem*

Vjerojatnosni prostor

σ - algebra je kolekcija \mathcal{F} podskupova skupa Ω , koja sadži \emptyset , Ω , komplement svakog svojeg člana $A^c = \Omega \setminus A$ i zatvorena je na prebrojivu uniju. Tada je zatvorena i na presjek.

Vjerojatnost je funkcija $p : \mathcal{F} \rightarrow [0, 1]$ za koju vrijedi:

- $p(A) \geq 0$
- $p(\Omega) = 1$
- $p(\Omega \setminus A) = p(A^c) = 1 - p(A)$
- $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

Vjerojatnosni prostor je trojka (p, \mathcal{F}, Ω) .

Vjerojatnost a'posteriori je statističke prirode.

Zadatak

*Prilikom 10000 pokušaja uspostave veze, ona uspostavljena 6789 puta.
Kolika je vjerojatnost uspostave veze?*

Konačan vjerojatnosni prostor

Primjer

Igra na sreću sastoji se od bacanja dviju kocaka. Odredite vjerojatnost da zbroj na kockama bude veći od 9? $\frac{6}{36}$.

Zadatak (Chevalier de Méré oko 1655.godine)

Koliko je vjerojatno da se u 4 bacanja jedne kocke bar jednom pojavi broj 6?

Zadatak

Kolika je vjerojatnost da se u 24 bacanja dviju kocaka barem jednom pojavi zbroj 12?

Prebrojiv vjerojatnosni prostor

Zadatak

Novčić se baca dok ne padne pismo. Kolika je vjerojatnost da pismo padne tek u petom pokušaju? $\frac{1}{32}$.

Zadatak

Kocka se baca dok se ne dobije 6. Koliko se puta mora bacati pa da s vjerojatnosti od 90% možemo tvrditi da će bar jednom pasti 6? 13, $p = 0,906536$.

Geometrijska vjerojatnost

Zadatak

Dečko i cura dogovore susret između 7 : 00 i 8 : 00 na trgu. Dogovore se, da ono koje dođe prvo, čeka drugo 20min. Kolika je vjerojatnost da će se ipak susresti?

Zadatak

Iz intervala $[0, 2]$ biraju se na slučajan način dva broja. Kolika je vjerojatnost da umnožak brojeva bude manji od 1?

Zadaci

- 1 Vennovim dijagramima pokažite $(A \cup B)^c = A^c \cap B^c$ i $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- 2 U studentskoj grupi su 24 studenta. Engleski govori 20 studenata, a njemački 10. Koliko ih govori oba jezika, ako dvojica ne govore ni njemački ni engleski? 8
- 3 Bacaju se tri novčića. Kolika je vjerojatnost da se na jednom pojavi pismo? 0.375
- 4 Bacaju se dvije kocke. Kolika je vjerojatnost da zbroj na kockama neće biti 9? 88.9%
- 5 U kvadrat je upisan krug. Kolika je vjerojatnost da pikado pogodi dio kvadrata izvan kruga, ako se pretpostavi da ne strijelac ne može promašiti metu? 21.56%
- 6 Na zidu dimenzija 8×5 metara nalaze se četiri prozora svaki dimenzija 1.8×1.2 metra. Kolika je vjerojatnost da dječak koji nasumice noću napucava loptu u zid pogodi bilo koji prozor? 21.6%

Uvjetna vjerojatnost

Vjerojatnost da se dogodi A ako se dogodio B :
$$p(A/B) = \frac{p(A \cap B)}{p(B)}.$$

Primjer

*Kocka je bačena i pao je neparan broj. Kolika je vjerojatnost da je pao tri?
Kolika je vjerojatnost da je pao broj dva?*

Primjer

Andrija, Blaž, Cvetko i Dragec kartaju belu. Andrija nema niti jednog herca. Kolika je vjerojatnost da Cvetko ima tri?

Zadatak

Mirko je zvao trefa iako u prvih šest mađarskih karata nije imao niti jednog. Kolika je vjerojatnost da u talonu obje karte budu trefovi?

Nezavisni događaji

Događaji A i B su nezavisni, ako informacija o događaju B ne utiče na vjerojatnost događaja A : $p(A/B) = p(A) = \frac{p(A \cap B)}{p(B)}$,
 $p(A \cap B) = p(A) \cdot p(B)$.

Primjer

Dobar, loš i zao strijelac gađaju vepra. Dobar strijelac na streljani pogađa u 80% gađanja, loš u 60% i zao u 90% gađanja. Koliko je vjerojatno da vepra pogodi samo jedan metak, ako sva trojica gađaju samo s po jednim metkom? Ako je vepra zaista pogodio jedan metak, koliko je vjerojatno da je metak ispalio dobar strijelac?

Zadaci

- 1 Tri strijelca gađaju jednu te istu metu. Vjerojatnost pogotka kod svakog strijelca redom iznosi 80%. Kolika je vjerojatnost
 - 1 točno jednog pogotka u metu 9.6%,
 - 2 bar jednog pogotka u metu 99.2%,
 - 3 najviše dva pogotka u metu 48.8%
- 2 Vjerojatnost da avion bude oboren prije nego što doprije do cilja je 5%. Ako doprije do cilja, avion ga uništava s vjerojatnosti od 40%. Koja je vjerojatnost da avion uništi cilj?(0,38)
- 3 Vjerojatnost da se otac dobije sina je 0.502. Koliko bi Miroslav trebao imati djece, pa da bude 99% siguran da će dobiti sina? 7

Potpuna vjerojatnost

Potpun sistem događaja - hipoteze $H_1, H_2, \dots, H_n \subset \Omega$, ako vrijedi:

$$H_1 \cup H_2 \cup \dots \cup H_n = \Omega \quad \text{i} \quad H_i \cap H_j = \emptyset$$

Formula potpune vjerojatnosti - $p(A) = \sum_{i=1}^n p(A/H_i) \cdot p(H_i)$.

Zadatak

Mercedes 60% proizvodnje ima u Turskoj, 30% u Kini i 10% u Njemačkoj. Postotak Mercedesova koji imaju kvar u garantnom roku, a napravljeni su u Turskoj iznosi 1%, za Kinu to je 1.5%, dok je za Mercedesove iz Njemačke 0.8%. Marko je kupio Mercedes. Kolika je vjerojatnost da će se pokvariti u garantnom roku?

Bayesova formula

Bayesova formula -

$$p(H_i/A) = \frac{p(H_i) \cdot p(A/H_i)}{p(A)}.$$

Zadatak

Kolika je vjerojatnost da je Markov Mercedes kineski?

Zadatak

Iz špila od 32 karte jedna je pala pod stol. Ako je Janko u prvih šest karata dobio tri trefa, kolika je vjerojatnost da pod stolom nije tref?

Zadaci

- 1 Strijelci Mate i Ante, svaki sa po jednim metkom, gađaju cilj. Mate pogađa u 70%, a Ante u 60% slučajeva. Utvrđeno je da je meta pogođena jednim metkom. Kolika je vjerojatnost da je metu pogodio Mate?
- 2 U uzorku ispitanika, u kojem je dio muškaraca 55%, 70% muškaraca i 60% žena su vozači. Kolika je vjerojatnost da slučajno odabrana osoba ima vozačku dozvolu? Kolika je vjerojatnost da je slučajno odabrani vozač muškarac?
- 3 U plavoj kutiji su tri bijele i dvije crne kuglice, a u crvenoj četiri bijele i pet crnih kuglica. Adam je iz crvene u plavu kutiju prebacio jednu kuglicu. Kolika je vjerojatnost da tada Martin iz plave kutije izvuče bijelu kuglicu? Ako je izvukao bijelu kuglicu, koliko je vjerojatno da je Adam baš crnu kuglicu prebacio prije izvlačenja?

Rješenja: 1. 61%; 2. 65.5% i 58.8%; 3. $\frac{26}{36}$ i $\frac{10}{26}$

Domaća zadaća za potpis

- 1 U neprozirnoj kutiji nalaze se 3 bijele i 4 crne kuglice. Na slučajan način izvlače se dvije kuglice. Kolika je vjerojatnost da su obje iste boje? (42.86%)
- 2 U poštanskom sustavu 45% paketa putuje vlakom, 40% cestovnim i 15% zračnim modom prijevoza. Udio inozemnih paketa ovisno o modovima je 26%, 35% i 22%. Kolika je vjerojatnost da slučajno odabrani paket u centrali bude inozemni? Ako netko dobije inozemni paket, koje su šanse da je doputovao avionom?
- 3 Student je pripremio 20 od 30 pitanja. Na ispitu se izvlače četiri i student smije samo jedno ne znati. Koliko je vjerojatno da neće pasti ispit?
- 4 Anketirano je 1200 vozača i 450 vozačica. Ustanovljeno je da 50%vozača nije imalo do sada niti jednu, a 35% samo jednu nesreću. Isto je ustanovljeno da 60% vozačica nije imalo niti jednu, a 30% samo jednu nesreću. Koliko je vjerojatno da osoba s više od jedne nesreće bude vozačica?
- 5 Na nekom fakultetu 75% studenata je položilo matematiku, 80% statistiku, a 60% oba ispita. Ako je student položio statistiku, koliko je vjerojatno položio i matematiku? Kolika je vjerojatnost da na fakultetu postoje studenti koji nisu položili nijedan navedeni predmet?

Zadaci uvjetne i Bayesove vjerojatnosti

- 1 Statistički podaci govore da autobus kasni u 10% slučajeva, a vlak u 35% slučajeva. Ako putnici u javnom prijevozu tri puta češće biraju autobus od vlaka,
 - 1 izračunajte kolika je vjerojatnost da slučajno odabrani putnik zakasni na određite?
 - 2 Ako znamo da je određeni putnik zakasnio, kolika je vjerojatnost da je išao vlakom?
- 2 Prilikom sudara u 5% slučajeva oba su vozača pod utjecajem alkohola i tada je vjerojatnost smrti 20%. U 25% je samo jedan vozač alkoholiziran, a vjerojatnost smrti je 15%. U 70% slučajeva oba su vozača trijezna i tada je vjerojatnost smrtnog ishoda 1%. Nakon sudara sa smrtnim posljedicama alkotestom je utvrđena alkoholiziranost jednog vozača. Kolika je vjerojatnost da je i onaj drugi pod utjecajem alkohola?

Rješenja: 1. 16.25% i 53.85%; 2. 21%

Zadaci s pismenih

- 1 Da bi osvojili iznos u TV-igri potrebno je točno odgovoriti na tri pitanja. Prvo pitanje zna otprilike 75% ljudi, drugo 50%, a treće 25%. Ako kandidat ne ponudi točan odgovor, igra se prekida. Kolika je vjerojatnost da
 - 1 TV kuća zadrži primamnjivi iznos? 71.875%
 - 2 kandidat ispadne za zadnjem pitanju? 28.125%
- 2 Na posljednjem roku student ima 4 ispita. Ako je prolaznost na ispitima 40%, kolika je vjerojatnost da student
 - 1 položi sve ispite? 2.56%
 - 2 jednog položi, a ostale ne položi? 34.56%
 - 3 da prvog položi, a ostale ne položi? 8.64%
- 3 Istraživanjem društvenih pojava dobiveno je da sindikat u 37% pregovora ostvari ciljeve. Ako se u sljedećih 10 godina očekuje 6 pregovora, kolika je vjerojatnost da sindikat
 - 1 ostvari sve svoje ciljeve? 0.256%
 - 2 polovično ostvari svoje ciljeve? 1.52%

Kombinatorni zadaci s pismenih

- 1 U četiri vagona na slučajan način ulazi 16 putnika. Znamo da u svakom vagonu ima točno 12 slobodnih mjesta. Kolika je vjerojatnost da
 - 1 u prvi vagon uđu četiri putnika?
 - 2 u svaki vagon uđe jednaki broj putnika?
 - 3 se neki od vagona popuni do posljednjeg mjesta?
- 2 Naplatne kućice autoputa imaju otvorena tri izlaza i svi se oni koriste jednako često. Ako vozači na slučajan način biraju koji će izlaz koristiti, kolika je vjerojatnost da
 - 1 od pet vozila njih tri izađe na istom izlazu?
 - 2 svih pet izađu kroz isti izlaz?

Razni zadaci s pismenih

- 1 Omjer putničkih, brzih i IC vlakova u željezničkom prometu je 5:3:2. Poznato je da 85% putničkih, 90% brzih i 95% IC vlakova dolazi na vrijeme. Ako je vlak u stanicu stigao na vrijeme, koliko je vjerojatno da je to IC vlak?
- 2 Poslodavac ima 25 zaposlenika koje može poslati na službeni put i od toga je 60% žena. Ako je na slučajan način odabran tim od petero ljudi, kolika je vjerojatnost da će u ekipi biti barem tri žene?
- 3 Od 20 poštarskih paketa 12 ih se mora isporučiti u prigradsko naselje. Ako poštar na slučajan način dobiva za isporuku 5 paketa, kolika je vjerojatnost da će barem jedan paket biti s isporukom u prigradsko naselje?

Razni zadaci s pismenih

- 1 U posljednjem tjednu prije ljetnih praznika student u pet dana izlazi na tri ispita. Ako su svi rasporedi ispita jednako vjerojatni, kolika je vjerojatnost da
 - 1 sve ispite polaže isti dan?
 - 2 u jednom danu polaže najviše jedan ispit?
 - 3 dva ispita polaže istog dana, a treći nekog drugog dana?
- 2 Vjerojatnost smanjene vidljivosti tijekom slijetanja zrakoplova iznosi 0.3. Odredite vjerojatnost da pri slijetanju:
 - 1 u nizu od 10 letova ne nastupi uvjet smanjene vidljivosti
 - 2 u nizu od 6 letova barem u polovici bude dobra vidljivost.

Slučajna varijabla. Uvod

Kakve su oklade moguće u sljedećim primjerima?

- Novčić se baca dok se ne dobije pismo.
- Broje se bacanja kocke dok se ne dobije šestica.
- U igri se bacaju dvije igraće kocke. Dobitak je zbroj na kockama.
- Oklada je na trenutak završetka nastave? Kakve mogu biti oklade?
- Ako se kladimo na masu koju će nam ujutro pokazati vaga, kakve su oklade moguće?

Zadatak

Koje su oklade zajedničke?

Definicija slučajne varijable

Definicija

- (Ω, \mathcal{F}, P) - vjerojatnosni prostor
- Slučajna varijabla X - funkcija $X : \Omega \rightarrow \mathbb{R}$
- obavezno - $P(X \leq x)$.

Zadatak

Definirati u prethodnim primjerima slučajne varijable.

Zadatak

Navedite po jedan primjer diskretne i jedan primjer kontinuirane slučajne varijable.

Funkcija razdiobe

Definicija

Neka je sve kao u prethodnom. Funkcija razdiobe $F : \mathbb{R} \rightarrow [0, 1]$ računa vjerojatnost: $F(x) = P(X \leq x)$.

Zadatak

U bacanju dviju kocaka igrač dobiva onoliko kuna koliko iznosi veći broj puta deset. Ako su brojevi isti, ne dobiva ništa. Napišite funkciju razdiobe.

Svojstva funkcije razdiobe:

- neopadajuća: $x_1 \leq x_2 \Rightarrow F(x_1) \leq F(x_2)$
- neprekidna sdesna $\lim_{\epsilon \rightarrow 0} F(x + \epsilon) = F(x)$
- vrijedi: $\lim_{x \rightarrow -\infty} F(x) = 0; \lim_{x \rightarrow \infty} F(x) = 1$.

Diskretna slučajna varijabla

Zadatak

U kutiji su tri crne i pet bijelih kuglica. Kuglice se mogu izvlačiti dok se ne izvuče bijela. Slučajna varijabla je broj izvlačenja.

- 1 *Koliko vrijednosti može imati slučajna varijabla?*
- 2 *Odredite pripadne vjerojatnosti za svaku vrijednost slučajne varijable.*

Definicija

- *Parovi (x_i, p_i) , $p_i = P(X = x_i)$ - distribucija vjerojatnosti*
- *Funkcija $f(x) = p(X = x)$ - gustoća vjerojatnosti*

Numeričke karakteristike

- Matematičko očekivanje $E(X) = x_1p_1 + x_2p_2 + \dots + x_np_n$
- Varijanca $V(X) = x_1^2p_1 + x_2^2p_2 + \dots + x_n^2p_n$
- Standardna devijacija $\sigma = \sqrt{V(X)}$.

Zadatak

Špil u šnapsu ima 20 karata od kojih je pet pikova. Igraču se dodijeli pet karata i po svakom piku ima pravo na 20 kuna. Koliki je očekivani dobitak, a kolika standardna devijacija?

Zadaci

Zadatak

Bežični modem pokušava uspostaviti vezu šest puta dnevno. Ako je vjerojatnost uspostave veze pri svakom pokušaju 70%, odredite očekivani broj uspješnih spajanja toga dana i standardnu devijaciju.

Zadatak

Igra na sreću sastoji se u bacanju četiri novčića jednog za drugim, ali se prekida ako padne glava. Za svaki bačeni novčić dobivaju se 2 kn, a za svaki koji nije bačen gube se 2 kn. Da li je igra poštena?

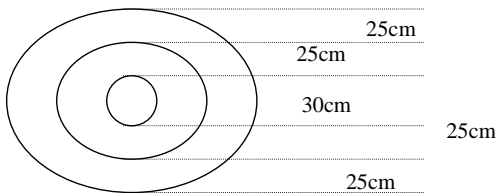
Zadatak

Igraču se iz špila od 32 dijeli 8. Ako po svakom hercu koji dobije uzima po 20 kn, koliko mora biti učešće igrača?

Pikado

Zadatak

Meta za pikado izgleda kao na slici. Središnji krug nosi 10 bodova, prvi srednji 5, a vanjski vijenac 1 bod. Igrač gađa metu s tri pikado strelice. Pod pretpostavkom da metu ne promašuje, izračunajte očekivani broj bodova, varijancu i standardnu devijaciju



Rješenja

- 1 $E(X) = 4.2$, $V(X) = 1.26$, $\sigma(X) = 1.12$
- 2 Nije, naklonjena je organizatoru, jer je $E(X) = -0.5$. Mogući ishodi su $-4, 0, 4$ i 8 kuna s vjerojatnostima redom 50% , 25% , 12.5% i 12.5% .
- 3 6.4 kune, a očekivani dobitci su od 0 preko $20, 40 \dots$ do 160 *kn*. Vjerojatnosti su redom 75% , 19.35% , $4.52\% \dots$ sve do 0.0000095% .
- 4 Za pikado $p(C) = \frac{9}{169}$; $p(S) = \frac{55}{169}$; $p(S) = \frac{109}{169}$. Dobici su redom $90, 65, 62, 40, 37, 34, 15, 12, 9$ i 6 *kn* s vjerojatnostima redom 0.02% , 0.28% , 0.53% , 1.69% , 6.46% , 6.17% , 3.45% , 19.74% , 37.69% i 23.98% . Očekivani dobitak je 13.40 *kn*, ali je standardna devijacija velika: $\sigma = 10.76$.

Binomna razdioba

Slučajan pokus s dva ishoda ponavlja se n puta:

- "uspjeh", s vjerojatnosti p
- "neuspjeh", s vjerojatnosti $q = 1 - p$

Slučajna varijabla X broji uspjehe: $X = \{0, 1, \dots, n\}$.

Funkcija (gustoće) vjerojatnosti -

$$f(x) = P(X = x) = \begin{cases} \binom{n}{x} \cdot p^x \cdot q^{n-x}, & x = 0, 1, \dots, n \\ 0, & x \neq 0, 1, \dots, n \end{cases}$$

Funkcija distribucije vjerojatnosti:

$$F(x) = P(X \leq x) = \sum_{k \leq x} \binom{n}{k} \cdot p^k \cdot q^{n-k}.$$

Matematičko očekivanje: $E(X) = np$

Varijanca: $V(X) = npq$

Zadatak

Zadatak

Marko na putu do posla prolazi kroz 5 nesinhroniziranih semafora na kojima zeleno traje $\frac{2}{5}$ faze. Na putovanju do posla broji semafore na kojima je imao prolaz. Koliki je očekivani broj prolaza? Koliki je očekivani broj stajanja? Kolika je varijanca prolaza? Kolika je varijanca stajanja? Kolika je vjerojatnost za najviše 3 prolaza? Kolika je vjerojatnost za bar tri prolaza?

Rješenja redom: 2,3,1.2,1.2,91.26%,31.74%.

Ispitni zadaci

Zadatak

*Vjerojatnost smanjene vidljivosti tijekom slijetanja zrakoplova je 30%.
Odredite vjerojatnost da pri slijetanju u nizu:*

- a) od 10 letova nastupi 3 uvjeta smanjene vidljivosti. 27%*
- b) od 6 letova dobra vidljivost bude u više od polovice slučajeva? 7%*

Zadatak

Vjerojatnost da se postigne dogovor vlade i opozicije je 0.25. Kolika je vjerojatnost da u četiri sastanka do dogovora

- uopće dođe. 68%*
- Koji je očekivani broj plodnih sastanaka i kolika je standardna devijacija? $E(X) = 1$, $\sigma = 0.866$*

Poissonova distribucija

- Zadana pozitivnim brojem λ
- X - broji rijetke događaje, $X = \{0, 1, 2, \dots\}$

Ako je λ očekivani broj rijetkih događaja, tada je:

$$p(X = x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}.$$

Funkcija distribucije, očekivanje, varijanca i standardna devijacija:

$$F(x) = \sum_{y \leq x} e^{-\lambda} \cdot \frac{\lambda^y}{y!}$$

$$E(X) = \lambda$$

$$V(X) = \lambda$$

$$\sigma = \sqrt{\lambda}$$

Ispitni zadaci

Zadatak

Otprilike 4% autobusa kasni. Na kolodvor dolazi dnevno 125 autobusa. Odredite vjerojatnost da tijekom dana

- 1 *kasni manje od 4 autobusa*
- 2 *kasni više od 3 autobusa*
- 3 *kasni samo jedan autobus.*

Zadatak

Broj vozila koja dolaze na servis tijekom jednog dana ima Poissonovu razdiobu. Prosječno u 20 radnih dana servis primi 45 vozila.

- 1 *Izračunajte standardnu devijaciju.*
- 2 *Kolika je vjerojatnost da u jednom danu bar jedno vozilo dođe na servis?*

Ispitni zadaci

Zadatak

Prosječni vremenski interval između dva poziva u centrali taxi službe je 20 sekundi. Ako se broj poziva u minuti ravna po Poissonovoj razdiobi, kolika je vjerojatnost da u minuti bude:

- 1 *barem 4 poziva 35%*
- 2 *samo jedan poziv? 15%*

Zadatak

Centrala taxi službe primi prosječno 120 poziva na sat, a mogu obraditi u minuti najviše 3 zahtjeva. Ako se broj poziva u minuti ravna po Poissonovoj razdiobi, kolika je vjerojatnost:

- 1 *preopterećenja centrale 14%*
- 2 *da u minuti bude samo jedan poziv? 27%*

Kontinuirana slučajna varijabla

Neka je (Ω, \mathcal{F}, P) vjerojatnosni prostor.

Slučajna varijabla $X : \Omega \rightarrow \mathbb{R}$ je kontinuirana

- ako vrijedi $P(X = x) = 0$
- ako postoji funkcija gustoće vjerojatnosti: $f : \mathbb{R} \rightarrow \mathbb{R}$:

Svojstva funkcije gustoće vjerojatnosti f :

$$(i) \int_{-\infty}^{+\infty} f(t) dt = 1$$

$$(ii) f(t) \geq 0 \text{ za svaki } t \in \mathbb{R}.$$

Primjer

*Odredite a^2 , tako da $f(t) = a^2 - t^2$ bude funkciju gustoće vjerojatnosti.
Rj. cca 0.9*

Funkcija distribucije

Ključna za računanje vjerojatnosti:

$$F(x) = \int_{-\infty}^x f(t)dt.$$

Moguće računanje

- $P(a < X < b) = P(a \leq X \leq b) = F(b) - F(a).$
- $P(X < b) = P(X \leq b) = F(b).$
- $P(X > a) = P(X \geq a) = 1 - F(a).$

Zadatak

Zapisati formulama funkciju distribucije za prethodni zadatak.

Numeričke karakteristike

Matematičko očekivanje $\mu = E(X) = \int_{-\infty}^{+\infty} tf(t)dt.$

Standardna devijacija $\sigma = \sqrt{V(X)}$, gdje je

$$V(X) = E [(X - E(X))^2] = E(X^2) - E(X)^2$$

Zadatak

Izračunati matematičko očekivanje i standardnu devijaciju za primjer prethodno zadane kontinuirane razdiobe gustoće vjerojatnosti.

Uniformna razdioba

Zadana je funkcijom gustoće vjerojatnosti

$$f(t) = \begin{cases} k, & a \leq t \leq b \\ 0, & \text{ostalo} \end{cases}$$

Zadatak

Izračunati parametar k ako je poznat interval $[a, b]$. Zapisati funkciju distribucije. Izračunati sve numeričke karakteristike uniformne razdiobe.

$$k = \frac{1}{b-a}, \quad F(x) = \frac{x-a}{b-a}, \quad \text{Za } a \leq x \leq b,$$

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}, \quad \sigma = \frac{b-a}{2\sqrt{3}} \quad \text{Za } x \leq a, \quad F(x) = 0, \quad \text{za } x \geq b, \\ F(x) = 1.$$

Eksponencijalna razdioba

Pozitivna konstanta $\lambda > 0$. Funkcija gustoće vjerojatnosti

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & \text{ostalo} \end{cases}$$

Funkcija distribucije: $F(x) = 1 - e^{-\lambda x}$ Matematičko očekivanje $E(X) = \frac{1}{\lambda}$

Zadatak

Prosječan dnevni promet na autocesti širokoj 30 m u špici sezone iznosi 24000 automobila. Medvjed trči 54 km/h. Koliko je vjerojatno da neće biti udaren onaj medo koji preskoči žicu i pretrčava autocestu?

Normalna razdioba

Normalna razdioba : $X \sim N(\mu, \sigma^2)$

Parametri : $\mu \in \mathbb{R}$ i $\sigma > 0$.

Funkcija gustoće vjerojatnosti :

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

Matematičko očekivanje : $EX = \mu$.

Varijanca : $V(X) = \sigma^2$.

Funkcija razdiobe :

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Standardna normalna razdioba

Standardna normalna razdioba : $Z \sim \mathcal{N}(0, 1)$.

Funkcija gustoće vjerojatnosti :

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Funkcija distribucije :

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du.$$

Tablice.

Uočiti : $\Phi(z < -3) = 0$ i $\Phi(z > 3) = 1$.

Supstitucija : $z = \frac{x - \mu}{\sigma}$ daje $F(x) = \Phi(z)$

Računanje vjerojatnosti :

$$p(a < x < b) = F(b) - F(a) = \Phi(b) - \Phi(a)$$

Primjena standardne normalne razdiobe

Zadatak

Odojci se kreću na sajmu cca 18 ± 4 kg. Odredite vjerojatnost da

- 1 odojak bude ispod 14 kg,*
- 2 odojak bude iznad 20 kg,*
- 3 odojak bude od 20 do 24 kg?*
- 4 Odredite gornju granicu za 80% odojaka.*
- 5 Odredite kilažu iznad koje je 3/4 odojaka.*

Kolokvij...

- 1 Broj vozila koja dolaze na servis ima Poissonovu razdiobu. Prosječno u roku 15 radnih dana servis primi 50 vozila.
 - 1 Izračunajte standardnu devijaciju.
 - 2 Kolika je vjerojatnost da u jednom danu barem jedno vozilo dođe na servis?
- 2 Prilikom ocjenjivanja nekog ispita koristi se normalna razdioba bodova na ispitu. Pozitivnu ocjenu dobiva gornjih 60% studenata na listi. Ako je na zadnjem roku očekivanje iznosilo 49 bodova, a standardna devijacija 16 bodova, koji je bio minimalni broj bodova za prolaznu ocjenu?

... s teškim zadatkom

1 Funkcija je zadana formulom $f(x) = \begin{cases} 0 & \text{za } x < 3 \\ a(x - 3) & \text{za } 3 \leq x < 5 \\ 0 & \text{za } x \geq 5 \end{cases}$

Odredite a tako da f bude funkcija gustoće vjerojatnosti slučajne varijable X . Nacrtati graf funkcije razdiobe $F(x)$. Izračunati $P(1 < X < 4)$, $P(4 < X < 6)$ i $P(X \geq 4)$.

... i nastavkom

- 1 Dva izlaza s autoputa koriste se jednako često. Kolika je vjerojatnost da od pet vozila
 - 1 tri vozila izađu na isti izlaz
 - 2 svi izađu kroz isti izlaz?
- 2 Ante pogodi vuka u 60%, Brane u 70%, a Čaruga u 90% slučajeva. Ako sva trojica zapucaju s po jednim metkom, a u vuku nađu dva metka, koliko je vjerojatno da Brane nije pogodio?

Kolokvij...

- ① Zlatko je iskovao novčić potpuno nalik običnom, ali koji na Pismo pada s vjerojatnosti duplo više nego na glavu. Imao ga, pa izgubio. Puno poslije, našao je sličan novčić i odlučio učiniti pokus:
 - a) ako novčić u 10 bacanja na Pismo padne barem 7 puta, Zlatko će ga zadržati (i varati s njim)
 - b) ako novčić padne manje od 7 puta, Zlatko će ga potrošiti zajedno s drugima.

Koliko je vjerojatno

- ① da Zlatko baci dugo traženi novčić?
- ② da Zlatko nadalje koristi običan novčić?

...s još dva...

- 1 U folklornom društvu na dva muška člana dolaze tri ženska. 30% muškaraca i 20% žena ima svjetliju kosu. Koliko je vjerojatno
 - 1 da slučajno odabrana osoba ima svjetliju kosu?
 - 2 da je izabrana osoba sa svjetlom kosom muškarac?
- 2 Dva čartera trebaju sletjeti između 18:00 i 19:00 sati. Koliko je vjerojatno da će između slijetanja proći barem pola sata?

...i završetkom

- 1 Tjelesna visina iznosi za žene 168cm sa standardnom devijacijom $\sigma = 6\text{cm}$. Odredite:
 - 1 postotak žena koje su više od 1.8m
 - 2 vjerojatnost da vaša šefica neće biti između 160 i 170 cm visine
 - 3 donju granicu ispod koje je 10% žena.
- 2 Zdenka uspješno parkira u 60% slučajeva. Koliko je vjerojatno da će u 5 pokušaja bar dvaput parkirati uspješno? Koliko je vjerojatno da će iz pet pokušaja najviše triput pogriješiti? Kolika bi bila vjerojatnost dobivanja oklade da će se Zdenki prilikom 5 uzastopnih parkiranja dogoditi nešto od navedenog? Pretpostavite binomnu razdiobu.
- 3 U kutiji je 10 proizvoda, od kojih su tri neispravna. Uzastopno se izvlače proizvodi dok se ne izvuče ispravan. Koliko je očekivanje broja izvučenih proizvoda, a kolika je standardna devijacija ovakovog izvlačenja?

Deskriptivna statistika

- Anketiranjem obitelji dobiveni su podaci o broju djece u obiteljima

1	2	1	3	1	0	1	1	1	1	2
3	4	1	3	4	1	1	1	1	0	2
1	0	2	1	2	5	1	6	0	1	3
2	5	4	4	3	1	1	4	3	2	5
4	4	3	1	1	1	0	3	2	2	4

- selektirajte, grupirajte, grafički prikažite i kvalitativno analizirajte podatke

Statistički skup. Modaliteti

- Skup statističkih jedinica
 - Pojmovna definicija: jednoznačna određenost elemenata
 - Prostorna definicija: pripadnost području
 - Vremenska definicija: trenutak ili interval prikupljanja
- Obilježje - svojstvo zajedničko statističkim jednicama po kojem se one razlikuju
- Modaliteti - stanja koja poprimaju obilježja.
- Populacija je skup svih podataka o promatranom svojstvu.

Primjer

Proizvodnja pšenice i prirod po 1 ha u tonoma u odabranim županijama Republike Hrvatske 2000. godine. (Statistički ljetopis, 2001. str 569 - 570)

Županija	Proizvodnja	Udio	Prirod	Zasijano
	<i>t</i>	%	<i>t/ha</i>	<i>ha</i>
Bjelovarsko - bilogorska	59 104		3.71	
Virovitičko - podravska	85 773		4.68	
Požeško - slavonska	56 495		4.85	
Brodsko - posavska	77 834		4.75	
Osječko - baranjska	103 623		5.12	
Vukovarsko - srijemska	199 847		5.13	
Ukupno				

Popunite tablicu, istaknite statistički skup i modalitete

Kvalitativno obilježje

Svojstvo po kojem su jedinice statističkog skupa slične ili se razlikuju.

- nominalno obilježje:
 - atributno, geografsko, alternativno
 - nisu dopuštene brojčane operacije
- ordinalno ili redosljedno obilježje
 - stručna sprema, ekonomska razvijenost
 - osim usporedbe, nikakve brojčane operacije nisu dozvoljene

Primjer

Živorodeni i umrli u odabranim općinama Hrvatskog Zagorja 1989. godine

Općina	Živorodeni	Umrli	Prirodni priraštaj	Živorodeni na 1 000 umrlih
1	2	3	4	5
Klanjec	123	188		
Krapina	319	342		
Pregrada	202	311		
Zabok	431	495		
Zlatar Bistrica	366	588		
Ukupno				

Popunite tablicu. Prikažite podatke grafički.

Kvantitativno obilježje

- Intervalno obilježje
 - temperatura, ocjena
 - sve računske operacije osim dijeljenja (omjera)
- Numeričko obilježje
 - neprekinuto ili kontinuirano - masa, vrijeme putovanja
 - diskretno - broj vozila koja ulaze u raskršće
 - sve računske operacije
 - diskretno obilježje s puno modaliteta - kontinuirano

Diskretna kontinuiranost

Dani su podaci o broju prekršaja koje su u mjesec dana evidentirali pripadnici patrolne prometne policije:

23	32	24	26	37	42	25	31	35	28	41	38
35	27	26	34	41	27	36	32	29	25	24	20
39	35	24	27	23	34	32	31	35	40	42	39

Što je statistički skup, a koji su modaliteti? Sortirajte podatke u razrede poštujući Sturgerovo pravilo o broju razreda:

$$k = 1 + \log_2 N,$$

gdje je N broj podataka.

Granični podatak svrstati uvijek u viši razred: $a \leq x_i < b$.

Sekundarni podaci

Tabela 1.8 Struktura zaposlenih u grani Proizvodnja obojenih metala u Republici Hrvatskoj u rujnu 1990 godine

Plaća u HRD	Struktura zaposlenih u %
<i>HRD</i>	P_i
1	2
4000.5-5000.5	8.8
5000.5-6000.5	17.1
6000.5-8000.5	50.7
8000.5-10000.5	20.0
10000.5-12000.5	3.0
12000.5-(20000.5)	0.4
Ukupno	

Napomena: U izvoru su dane nominalne granice razreda.

Izvor: Statistički godišnjak Republike Hrvatske, 1991

- Izračunajte prosječnu plaću.

Centralna tendencija

- **Medijan** je vrijednost numeričkog obilježja koje dijeli niz numeričkih podataka poredanih po veličini na dva jednaka dijela.
- **Kvartili** dijele niz na četiri jednaka dijela.
- Interkvartil čine podaci između prvog i trećeg kvartila.
- **Percentili** na 100. **Fraktili** su percentili izraženi decimalnim brojevima.
- **Mod** je vrijednost obilježja koje se najčešće javlja.

Grupiranje

Iznosi računa na blagajni samoposluge: (kn)

620	180	510	420	190	430	170	200	100
390	130	540	160	230	260	470	570	820
520	180	870	240	170	150	380	260	480
290	490	150	250	510	710	320	360	130
590	680	160	270	190	240	520	170	230

Zadatak

Poredajte podatke po veličini i odredite medijan i mod. Svrstajte podatke u razrede s pravim granicama i izračunajte aritmetičku sredinu i standardnu devijaciju.

Inferencijalna statistika

- **Apsolutna frekvencija:** f_i intenzitet pojavnosti $\sum f_i = N$.
- **Kumulativna frekvencija:** F_i , zbraja pojavnosti do uključivo i -tog modaliteta $F_n = N$
- **Relativna frekvencija** P_i, p_i ili f_{ir} , udio: $p_i = \frac{f_i}{N}$.

Aritmetička sredina

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

Standardna devijacija

$$\sigma^2(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i - 1} = \frac{N}{N-1} \sum_{i=1}^n (x_i - \bar{x})^2 p_i.$$

- **Statistički niz:** parovi modalitet-frekvencija

Primjer

Potrošnja goriva kamiona u voznom parku iskazana u $l/100km$ donosi podatke

14	12	15	11	11	12	15	13	13	13	11	16	17
12	13	12	11	12	13	15	14	13	14	15	17	15
12	12	14	11	13	13	12	14	13	12	15	17	16

Selektirajte podatke u statistički niz, navedite apsolutne, kumulativne, relativne i kumulativne relativne frekvencije. Izračunajte aritmetičku sredinu i standardnu devijaciju sortiranih podataka.

Zadatak

U mjesečnom statističkom izvješću DSZ, br. 4, 2001, na str. 92. nalaze se podaci o mirovinama u Republici Hrvatskoj, stanje potkraj travnja 2001. Podaci se odnose na umirovljenike bez zaštitnog dodatka. Mirovina u *kn*, broj korisnika u (000),

Mirovina	Broj korisnika	Udio	Kumulativno
x_i	f_i	p_i	F_i
(300)-500	95.4		
500-1000	215.3		
1000-1500	254.3		
1500-2000	219.3		
2000-4000	180.4		
4000-8000	12.4		
Ukupno			

Popunite tablicu. Izračunajte prosječnu mirovinu. Izračunajte standardnu devijaciju prosječne mirovine. Odredite interkvartil

Uvod u teoriju uzoraka

- 1 Populacija - očekivanje μ i varijanca σ^2 (nepoznato)
- 2 Svaki podatak iz populacije odabran slučajno: (X_1, X_2, \dots, X_n)
- 3 Svaka slučajna varijabla X_i - očekivanje μ i varijanca σ^2
- 4 Statistika - slučajna varijabla izračunata iz uzorka
- 5 Aritmetička sredina uzorka $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$
- 6 Varijanca uzorka $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \approx \sigma^2$.

Teorem

Uz prethodno definirane pojmove

$$E(\bar{X}) = \mu, \quad E(S) = \sigma.$$

Statistička procjena aritmetičke sredine

Uzorak treba imati barem 30 elemenata populacije: $n \geq 30$.

- Varijanca aritmetičke sredine $V(\bar{X}) = \frac{n\sigma^2}{n^2} \approx \frac{S^2}{n}$.
- Statistika aritmetičkih sredina $\bar{X} \sim \mathcal{N}\left(\mu, \frac{S^2}{n}\right)$
- Tablično: $\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$

Primjer

Duljina putovanja na posao izmjerena na 100 građana iznosila je 15 ± 2 minute. Izračunajte interval stvarnog putovanja na posao koji vrijedi u 95% slučajeva. Odredite donju granicu putovanja u 90% slučajeva. Odredite gornju granicu putovanja u 99.9% slučajeva.

Rješenje prvog zahtjeva

- Uzorak: $n = 100$, $\bar{X} = 15$, $S = 2$.
- Rješavati jednadžbu $c = ?$, stvarni μ nepoznat i koristiti tabličnu razdiobu

$$P(\bar{X} - c < \mu < \bar{X} + c) = 95\%$$

$$P\left(-\frac{c}{\frac{S}{\sqrt{n}}} < \frac{\mu - \bar{X}}{\frac{S}{\sqrt{n}}} < \frac{c}{\frac{S}{\sqrt{n}}}\right) = 0.95$$

$$D\left(\frac{c}{\frac{S}{\sqrt{n}}}\right) = 0.95$$

$$c = 1.96 \frac{2}{10} = 0.392$$

- Traženi interval je od 14.6 do 15.4 min. U 95% anketa to će biti prosjek.

Rješenje drugog zahtjeva

- Uzorak: $n = 100$, $\bar{X} = 15$, $S = 2$.
- Rješavati jednadžbu $c = ?$, stvarni μ nepoznat i koristiti tabličnu razdiobu

$$\begin{aligned}P(\bar{X} - c < \mu) &= 90\% \\P\left(\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} < \frac{c}{\frac{S}{\sqrt{n}}}\right) &= 0.9 \\ \Phi\left(\frac{c}{\frac{S}{\sqrt{n}}}\right) &= 0.9000 \sim 0.8997 \\ c &= 1.28 \frac{2}{10} = 0.256\end{aligned}$$

- U 90% anketa će najmanja vrijednost prosjeka biti najmanje 14.7 min.

Rješenje trećeg zahtjeva

- Uzorak: $n = 100$, $\bar{X} = 15$, $S = 2$.
- Rješavati jednadžbu $c = ?$, stvarni μ nepoznat i koristiti tabličnu razdiobu

$$\begin{aligned}P(\mu < \bar{X} + c) &= 99.9\% \\P\left(\frac{-c}{\frac{S}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}\right) &= 0.999 \\ \Phi\left(\frac{-c}{\frac{S}{\sqrt{n}}}\right) &= 0.001 \sim 0.0013 \\ c &= 3 \cdot \frac{2}{10} = 0.6\end{aligned}$$

- U 99.9% anketa će najveća vrijednost prosjeka biti manja od 15.6 min.

Zadaci

Zadatak

Odredite interval mogućeg putovanja na posao koji bi vrijedio u 60% slučajeva?

Rješenje $14.8 \leq \bar{x} \leq 15.2$

Zadatak

Broj popušenih cigareta dnevno kod teenagerskih pljugera nakon ankete dan je u tablici

<i>teenageri</i>	30	21	14	13	8	3	3	2	4	2
<i>cigarete</i>	1	2	3	4	5	6	7	8	9	10

Koliko prosječno cigareta jedan tinejder popuši dnevno? Odredite interval popušenih cigareta za 50% teenpljugera. Odredite gornju granicu popušenog iznad koje puši samo 1% pušača.

Sirovi uzorak

Zadatak

Uzorak se sastoji od podataka o prodaji goriva (supera95) po putničkom automobilu na benzinskoj postaji.

35	24	19	10	11	16	9	33	38	29	7	7
24	5	33	33	11	15	17	11	11	22	7	15
39	14	17	27	21	28	35	24	26	20	28	9
14	31	31	5	18	11	25	15	37	36	11	23
10	34	15	14	25	23	25	21	36	13	6	27

Procijenite prosječnu prodaju benzina po automobilu. Odredite interval prosječne prodaje benzina po automobilu za 75% potrošača. Ispod koje granice troši 10% potrošača?

Procjena udjela u populaciji

- Neka je p nepoznati postotak odabranih u populaciji
- Neka je n veličina uzorka iz populacije
- Neka je Y broj odabranih u uzorku
- Neka je $\bar{p} = \frac{Y}{n}$ postotak odabranih u uzorku.

- Tada je
$$\frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx \frac{\bar{p} - p}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}} \sim \mathcal{N}(0, 1),$$

jer je varijanca udjela po uzorcima $S^2 = \bar{p}(1 - \bar{p}) \sim p(1 - p) = \sigma^2$.

Anketa

Primjer

Ispitivanjem javnog mnijenja utvrđeno je da ulazak u EU opravdava 44% anketiranih. Ispitati standardnu devijaciju i odrediti interval postotka uz vjerojatnost pogreške 5%, ako je broj ispitanih bio

- 1 130
- 2 1300
- 3 13000
- 4 13

Rješenje

Analogno, $P(\bar{p} - c < p < \bar{p} + c) = 95\% \sim 0.9749$ daje redom granice unutar $44\% \pm 8.5\%$, 2.7% , 0.85% i $\pm 27\%$.

Testiranje statističkih hipoteza

Dobra pretpostavka - nulta hipoteza H_0 . Alternativa je pogrešna - H_1 .

- 1 Prihvatiti pogrešnu pretpostavku - manja greška - druge vrste vjerojatnost se naziva β
- 2 Odbaciti dobru pretpostavku - veća greška - prve vrste vjerojatnost se naziva α

Primjer

Zlatko je imao novčić koji je triput češće padao na pismo nego na glavu i izgubio ga. Nakon što je našao sličan novčić, testirao ga je: ako padne barem 7 puta na pismo u 10 bacanja, to je pravi. Ako nije, nikom ništa. Kolika je pogreška prve, a koliko pogreška druge vrste?

$$\alpha = 22.4\%, \beta = 17.2\%$$

Konstrukcija testa

- 1 dopuštena vjerojatnost greške prve vrste - α - razina signifikantnosti
- 2 minimalna vjerojatnost greške druge vrste - β
- 3 pouzdanost testa: $1 - \alpha$
- 4 jakost ili oštrina testa: $1 - \beta$

Primjer

Trajanje ekrana je 1000 ± 200 sati. Tehnolozi tvrde da novi ekrani traju dulje. Konstruirajte test na 25 novih televizora koji bi potvrdio dulje trajanje uz 2% signifikantnosti.

Rješenje

Ustanoviti \bar{X} iznad kojeg svaki dobiveni prosjek opovrgava nultu hipotezu $H_0: \mu = 1000, \sigma = 200$, a u korist alternative $H_1: \mu > 1000$.

Detalji rješenja i napomene

- Zahtjevani \bar{X} dobiva se rješavanjem jednadžbe po c iz uvjeta

$$P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > c\right) = 2\% \sim 0.02$$

- Iz $\Phi(c) = 0.9798$ slijedi $c = 2.05$, pa je

$$\bar{X} > 1000 + 2.05 \cdot \frac{200}{\sqrt{25}} = 1082.$$

- Vjerojatnost $\beta = P(\bar{X} \leq 1082)$, uz uvjet $\mu > 1000$ je nemoguće utvrditi radi nepoznatog μ .

Zadatak

Utvrđite vjerojatnost da Vam test ne prolazi, ako zaista novi ekrani traju 100 sati dulje, no rasipanje rezultata trajana je 40%?

Hipoteze o vjerojatnosti

Testira se $Z(p) = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathcal{N}(0, 1)$. Testiranje alternativa

$$\frac{H_0 : p = p_0}{H_1 : p \neq p_0} \parallel \frac{H_0 : p = p_0}{H_1 : p < p_0} \parallel \frac{H_0 : p = p_0}{H_1 : p > p_0}$$

ruši H_0 uz pouzdanost α ako je

$$\begin{array}{l} P(Z(p) < c_1) = \frac{\alpha}{2} \\ P(Z(p) > c_2) = \frac{\alpha}{2} \end{array} \parallel P(Z(p) < c) = \alpha \parallel P(Z(p) > c) = \alpha$$

Napomena

U praksi se zada α , odrede se kritične vrijednosti, a nakon toga se računa β . Ako je β prevelik, uzme se veći uzorak

Primjer

Primjer

Čelnici stranke A tvrde da njihovu stranku podupire barem 22% birača. Ispitivanje se provodi na uzorku od 1 000 ljudi.

- 1 Formirajte test koji će s greškom 1. vrste $\alpha = 0.05$ odbacivati ili prihvaćati gornju tvrdnju.
- 2 Kolika je vjerojatnost da će Vaš test prihvatiti gornju tvrdnju ako u stvarnosti tek 18% birača podržava stranku A?

Manje od 199 ljudi od nasumce odabranih 1 000 daje glas za. Anketa prihvaća u 6% slučajeva.

Zadatak

Zadatak

Čelnici jedne stranke tvrde da vladu podupire najviše 40% birača. Ispitivanje se provodi na uzorku od 500 ljudi.

- 1 Formirajte test koji će s greškom prve vrste $\alpha = 2\%$ prihvatiti ili odbaciti gornju tvrdnju.
- 2 Kolika je vjerojatnost da će vaš test prihvatiti gornju tvrdnju ako u stvarnosti 50% birača podržava vladu?

Rješenje

- 1 Iznad 223 od 500 ispitanika odbacuje tvrdnju opozicije
- 2 vjerojatnost je 0.7%

Zadatak

Zadatak

Vlada tvrdi da je protiv ulaska Hrvatske u EU manje od 50% građana.

- 1 *Formirajte test na uzorku od 400 ljudi koji će uz grešku prve vrste $\alpha = 0.05$ prihvatiti ili odbaciti tvrdnju Vlade.*
- 2 *Kolika je vjerojatnost da će Vaš test prihvatiti tvrdnju Vlade ako je u stvarnosti 55% građana protiv?*

Rješenje

- 1 *Ako više od 208 od njih 400 bude protiv, vlada ne bi trebala tako nešto pričati!.*
- 2 *Šanse su 11.9%!.*

χ^2 test

Testiranje podudarnosti empirijskih frekvencija f_i i teoretskih frekvencija f_i^* veličinom

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - f_i^*)^2}{f_i^*}.$$

Hipoteza o podudarnosti se odbacuje ako je $\chi^2 > c$ sa vjerojatnosti

$$F(c) = 1 - \alpha$$

da se frekvencije zaista ne podudaraju.

Testiranje kontinuiranih razdioba

Zadatak

Mjerenjem vijeka trajanja akumulatora u mjesecima za 5000 transformatora dobiveni su sljedeći rezultati.

trajanje	0-10	10-20	20-30	30-40
transformatori	2250	930	370	140

Uz razinu pouzdanosti 1% testirajte hipotezu da je vijek trajanja transformatora veličina eksponencijalne distribucije.

Zadatak

Dani su podaci pružatelja internet usluga o sekundama trajanja prekida

veze (linka):	sekunde	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5
	prekidi	4	83	281	79	3

Testirajte hipotezu o normalnoj razdiobi uz nivo signifikantnosti 5%.

Testiranje diskretnih razdioba

Zadatak

Broj studenata koji su na ispitu točno riješili odgovarajući broj zadataka prikazan je u tablici:

broj riješenih zadataka	0	1	2	3	4
broj studenata	8	10	12	4	2

Testirajte hipotezu o Poissonovoj razdiobi broja točno riješenih zadataka na testu, ako je nivo signifikantnosti $\alpha = 5\%$.

Zadatak

Prilikom anketiranja studenata o broju izlazaka na ispit iz vjerojatnosti i statistike dobiveni su slijedeći odgovori:

izlasci	0	1	2	3	4
studenata	6	12	18	16	7

Testirajte hipotezu o binomnoj razdiobi broja izlazaka na ispit. Nivo signifikantnosti neka je 5%.

Korelacija i regresija

Koeficijent korelacije dvaju obilježja mjerenih na istih n statističkih jedinica jednak je

$$r = \frac{S_{xy}}{S_x S_y},$$

gdje je

$$S_{xy} = \frac{1}{n-1} \sum_i (x_i - \bar{x})(y_i - \bar{y}), \text{ kovarijanca}$$

$$S_x^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$$

$$S_y^2 = \frac{1}{n-1} \sum_i (y_i - \bar{y})^2$$

Koeficijent operativno

Razrađena formula za $\begin{array}{c|ccccc} x & x_1 & x_2 & \cdots & x_n \\ y & y_1 & y_2 & \cdots & y_n \end{array}$ glasi

$$r = \frac{\sum_i x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum_i x_i^2 - n\bar{x}^2) \cdot (\sum_i y_i^2 - n\bar{y}^2)}}.$$

Zadatak

Odredite koeficijent korelacije potrošnje vina i pive kroz godine u milijunima litara

<i>god</i>	1995.	1996.	1997.	1998.	1999.
<i>vino</i>	13.3	12.2	11.8	10.7	10.5
<i>pivo</i>	2.9	2.7	2.7	2.4	2.4

Pravac regresije za uzorak

U slučaju apsolutno velikih n -ova, isplati se pisati pravac regresije $y = a + bx$ o ovisnosti y -a o x -u:

$$a = \frac{\sum_i y_i \cdot \sum_i x_i^2 - \sum_i x_i \cdot \sum_i x_i y_i}{n \cdot \sum_i x_i^2 - (\sum_i x_i)^2}$$
$$b = \frac{n \cdot \sum_i x_i y_i - \sum_i x_i \cdot \sum_i y_i}{n \cdot \sum_i x_i^2 - (\sum_i x_i)^2}$$

Zadatak

Odredite jednadžbu pravca regresije ovisnosti potrošnje piva o potrošnji vina.

Korelacija s višestrukim parovima podataka

Ako parovi imaju višestruke frekvencije, zapis se komplicira:

x^y	y_1	y_2	\cdots	y_n
x_1	f_{11}	f_{12}	\cdots	f_{1n}
x_2	f_{21}	f_{11}	\cdots	f_{11}
\vdots	\vdots	\vdots	\vdots	\vdots
x_m	f_{m1}	f_{m1}	\cdots	f_{mn}

Formula nije jednostavna

$$r = \frac{n \cdot \sum \sum x_i y_j f_{ij} - (\sum x_i f_i)(\sum y_j f_j)}{\sqrt{[n \cdot \sum x_i^2 f_i - (\sum x_i f_i)^2][n \cdot \sum y_j^2 f_j - (\sum y_j f_j)^2]}}$$

Zadatak

Izračunajte koeficijent korelacije između plaće u tisućama kuna i broja automobila koje su razbili za nekoliko ispitanih osoba.

plaća \ auti	5	4	3	2
4			2	3
8		5	2	2
12	4	3		

Dopunski zadaci

- 1 Broj zaposlenih i ukupan prihod zadani su u tablici:

zaposleni	22	31	90	82	43
prihod	250	300	920	850	410

Izračunajte koeficijent korelacije i napišite pravac regresije po kojem prihod ovisi o broju zaposlenih.

- 2 Testirajte hipotezu Poissonove distribucije uz nivo signifikantnosti 0.05 za kontrolu 200 serija nekog proizvoda:

neispravnih u seriji	0	1	2	3	4
broj serija	112	54	28	4	2

- 3 Mjerenjem vijeka trajanja akumulatora u godinama dobiveni su podaci:

trajanje	0-1	2-3	3-4	4-5	5-6
broj akumulatora	150	100	70	45	25

Testirajte hipotezu o eksponencijalnoj distribuciji.

Rješenja dopunskih zadataka

- ① $y = 10.285x - 5.28$
- ② Odbacuje se uz $\chi^2 = 6.04$. $\bar{x} = 0.65$
- ③ $\bar{x} = 2$, $\lambda = 0.5$, $\chi^2 = 111.508 > 7.81$.

Ispit iz vjerojatnosti ...

- 1 U plavoj kutiji su četiri crne i tri bijele kuglice, a u žutoj kutiji su tri crne i četiri bijele kuglice. Martin je iz plave u žutu kutiju prebacio jednu kuglicu. Koliko je vjerojatno da će Adam iz žute kutije nakon toga izvući crnu kuglicu?
- 2 Mjesečni broj prometnih nesreća u Republici Hrvatskoj slučajna je varijabla normalne razdiobe s očekivanim brojem od 350 i standardnim odstupanjem od 70 prometnih nesreća.
 - 1 Koliko je vjerojatno da u lipnju bude najviše 400 prometnih nesreća?
 - 2 Do danas je zabilježeno 150 prometnih nesreća. Koliko je vjerojatno da ih do kraja mjeseca neće biti više od 150?

... i statistike

- 1 Broj neispravnih autobusa po danima zabilježen je u tablici:

neispr. bus.	0	1	2	3
dani	10	12	6	2

Testirajte hipotezu o Poissonovoj razdiobi uz signifikantnost $\alpha = 1\%$.
Koliko je vjerojatno da će sutra bar jedan autobus biti neispravan?

- 2 Izračunajte koeficijent korelacije za visinu mladenca i mladenke u sljedećim vjenčanjima, a u *cm*:

mladenka	162	164	158	170	160
mladenac	172	175	178	168	182

Nacrtajte pravac regresije.