VJEROJATNOST

⇒ SLUČAJNI POKUS ⇒ svaki pokus kojem ne znamo ishod

⇒ ELEMENTARNI DOGAĐAJ ⇒ osnovni ishod pokusa

↳ PR: *bacanje kocke:*

DOGAĐAJ: pao je paran broj

ELEMENTARNI DOGAĐAJ: pala je dvojka

⇒ SKUP SVIH ELEMENTARNIH ISHODA:

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

⇒ nemoguć događaj (prazan skup) - oznaka $\emptyset$

⇒ DOGAĐAJ je stoga svaki podskup od Ω

⇒ De Morganovi zakoni:

a) $\overline{A \cup B} = \bar{A} \cap \bar{B}$

b) $\overline{A \cap B} = \bar{A} \cup \bar{B}$

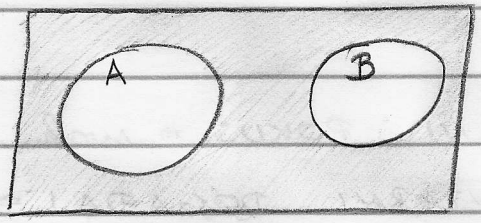
↳ Dokaz a): $\omega \in \overline{A \cup B} \Leftrightarrow \omega \notin A \cup B \Leftrightarrow$

$\Leftrightarrow \omega \notin A$ i $\omega \notin B \Leftrightarrow$

$\Leftrightarrow \omega \in \bar{A}$ i $\omega \in \bar{B} \Leftrightarrow \omega \in \bar{A} \cap \bar{B}$

↳ analogan dokaz tvrdnje pod „b"

⇒ događaji A i B su **disjunktni**, ako vrijedi da je $A \cap B = \emptyset$



⇒ ALGEBRA DOGAĐAJA:

↳ to je svaka familja $\mathcal{F}$ podskupova od Ω na kojoj je definirana binarna operacija „+" i operacija komplementa sa svojstvima:

1) $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$
2) $A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$
3) $A, B \in \mathcal{F} \Rightarrow A + B \in \mathcal{F}$

↳ NAPOMENA: operator „+" je ustvari unija skupova „$\cup$"

↳ primjer da presjek slijedi iz trećeg svojstva:

$$A \cdot B \Rightarrow \overline{\overline{A} + \overline{B}} \in \mathcal{F}$$

DEFINICIJA VJEROJATNOSTI

⇒ VJEROJATNOST je preslikavanje (funkcija) takvo da $P: \mathcal{F} \to [0,1]$ definirano na algebri $\mathcal{F}$ sa svojstvima:

1) NORMIRANOST: $P(\Omega) = 1$

$P(\emptyset) = 0$

2) MONOTONOST: ako imamo $A \subseteq B$, tada vrijedi $P(A) \leq P(B)$

3) ADITIVNOST: ako su A i B disjunktni, tada je:

$$P(A \cup B) = P(A) + P(B)$$

⇒ iz gore navedenih svojstava, slijede neki rahljući:

- vjerojatnost komplementa: $P(\bar{A}) = 1 - P(A)$

  ↳ dokaz:

$$1 = P(\Omega) = P(A \cap \bar{A}) = P(A) + P(\bar{A})$$

(jer su A i $\bar{A}$ disjunktni)

- vjerojatnost unije: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

ZAD: Ako neko preslikavanje ima svojstva $P(\Omega) = 1$ i $P(A \cup B) = P(A) + P(B)$ za disjunktne skupove, dokažite da je tada P vjerojatnost!

↳ stoga, mi trebamo dokazat monotonost i svojstvo da je $P(\emptyset) = 0$

i) $P(\emptyset) = 0$

$$P(\emptyset \cup \Omega) = P(\emptyset) + P(\Omega) = P(\emptyset) + 1$$

$$\downarrow \quad P(\emptyset \cup \Omega) = 1$$

$$1 = P(\emptyset) + 1 \quad \Rightarrow \quad P(\emptyset) = 0$$

ii) monotonost : $A \subseteq B$

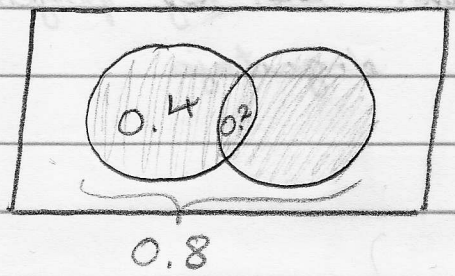
$$P(B) = P\big(A \cup (B \setminus A)\big) = P(A) + \underbrace{P(A \setminus B)}_{\geq 0}$$

↳ iz gornjega slijedi :

$$P(B) \geq P(A)$$

ŽAD: Neka je $P(A \cup B) = 0.8$, $P(AB) = 0.2$, $P(\bar{A}) = 0.6$. Izračunaj sljedeće:

a) $P(B) = ?$



$P(A \cup B) = P(A) + P(B) - P(AB)$

$P(B) = 0.6$

b) $P(\bar{A}\bar{B}) = 1 - P(A \cup B) = 1 - 0.8 = 0.2$

c) $P(A\bar{B}) = P(A) - P(AB) = 0.2$

KONAČNI VJEROJATNOSNI PROSTOR

⇒ to je vjerojatnosni prostor Ω koj posjeduje konačno mnogo elementarnih događaja

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

$$\left.\begin{aligned} P_1 &= P(\omega_1) \\ P_2 &= P(\omega_2) \\ &\vdots \\ P_n &= P(\omega_n) \end{aligned}\right\} \sum P_i = 1$$

PR: D'Alembertov problem: bacanje dva novčića!

↳ mogući ishodi:

$\omega_1 = \{2 \text{ pisma}\}$

$\omega_2 = \{2 \text{ glave}\}$

$\omega_3 = \{1 \text{ pismo}, 1 \text{ glava}\}$

} KRIVO! jer je D'Alembert zaključio $P_1 = P_2 = P_3 = \frac{1}{3}$ što je krivo

↳ D'Alembertov zaključak je bio pogrešan jer je krivo portovio elementarne događaje

↳ prvi elementarni ishodi su ovakvi:

$\omega_1 = \{$ na prvom pismo, na drugom pismo$\}$

$\omega_2 = \{$ na prvom glavo, na drugom glave$\}$

$\omega_3 = \{$ na prvom glavo, na drugom pismo$\}$

$\omega_4 = \{$ na prvom pismo, na drugom glava$\}$

↳ sada su svi elementarni događaji jednako vjerojatni

↳ NAPOMENA: prinjet da je redoslijed bacanja novčića bitan

⇒ ako su svi elementarni događaji jednako vjerojatni, tada vrijedi:

$$\sum P_i = N \cdot P_i = 1 \quad \Rightarrow \quad \boxed{P_i = \frac{1}{N}}$$

↳ ovakav vjerojatnosni prostor naziva se KLASIČNI vjerojatnosni prostor

↳ tada je vjerojatnost događaja A:

$$\boxed{P(A) = \frac{M}{N} = \frac{\text{broj povoljnih ishoda}}{\text{ukupan broj ishoda}}}$$

⇒ dva bitna pravila za radatke:

a) PRODUKTNO PRAVILO

b) ODABIR „k" ELEMENATA OD NJIH „n" ⇒ $\binom{n}{k}$

ZAD: Bacili smo 4 kocke. Izračunaj vjerojatnost:

a) dobili smo različite brojeve

b) dobili smo točno dvije šestice

c) dobili smo barem dvije šestice

d) točno 3 broja manja od 5

e) dva (različita) para jednakih brojeva

a) $P(A) = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4}$

b) $P(B) = \frac{\binom{4}{2} \cdot 5^2}{6^4}$

c) $P(C) = \frac{\binom{4}{2} \cdot 5^2 + \binom{4}{3} \cdot 5 + \binom{4}{4}}{6^4}$

d) $P(D) = \frac{\binom{4}{3} \cdot 4^3 \cdot 2}{6^4}$

e) $P(E) = \frac{\binom{6}{2}\binom{4}{2}}{6^4}$

ZAD: Izvlačimo 5 karata od 52 karte (imamo 4 „boje") (imamo 13 vrijednosti). Izračunaj vjerojatnosti za:

a) flush (5 istih boja)

b) royal flush (skala od 10 do asa u boji)

c) full house (3 karte iste jačine i 2 neke druge)

d) two pair (dva para)

e) one pair (jedan par)

f) straight (skala od 5 karata da nisu iste boje)

a) $P(A) = \frac{\binom{4}{1}\binom{13}{5}}{\binom{52}{5}}$

b) $P(B) = \frac{4}{\binom{52}{5}}$

c) $P(C) = \frac{13 \cdot \binom{4}{3} \cdot 12\binom{4}{2}}{\binom{52}{5}}$ } pripazit da nebi bilo dobro $\binom{13}{2}$ jer bi onda izračunali 3 PIKA + 2 KARE i 2 PIKA + 3 KARE

d) $P(D) = \frac{\binom{13}{2}\binom{4}{2}\binom{4}{2} \cdot 11 \cdot 4}{\binom{52}{5}}$

e) $P(E) = \frac{\binom{13}{1}\binom{4}{2} \cdot \binom{12}{3} 4^3}{\binom{52}{5}}$ } pripazit da nebi bilo dobro $13 \cdot 12 \cdot 11$ jer bi onda fiksirali redoslijed

f) $P(F) = \frac{\binom{10}{1}\left(4^5 - 4\right)}{\binom{52}{5}}$

odabir skale

na svaku kartu biramo boju, ali moramo oduzeti slučaj da su sve karte iste boje

ZAD: Između 7 bračnih parova biramo 5 osoba. Izračunaj vjerojatnost:

a) svih 5 osoba su istog spola?

b) nije jedan bračni par među njih 5?

a) $P(A) = \frac{2 \cdot \binom{7}{5}}{\binom{14}{5}}$

b) $P(B) = \frac{\binom{7}{5} 2^5}{\binom{14}{5}}$

$\binom{7}{5}$ → biramo bračne parove

2^5 → biramo osobu iz bračnog para

ZAD: 12 putnika ulazi u 4 vagona. Izračunaj vjerojatnost da su:

a) u prvi i treći vagon ušla po 3 putnika ?

b) u svaki vagon ušla po 3 putnika

a) $$P(A) = \frac{\binom{12}{3} \cdot \binom{9}{3} \cdot 2^6}{4^{12}}$$

b) $$P(B) = \frac{\binom{12}{3}\binom{9}{3}\binom{6}{3}}{4^{12}}$$

PONAVLJANJE POKUSA

PR: Bacamo kocku 3 puta. Izračunaj vjerojatnost da smo tri puta dobil šesticu?

$$P(A) = \underbrace{\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}}_{} =$$

tri neovisna događaja, pa te vjerojatnost možemo množiti //

↳ primjet da smo u ovom slučaju ovo mogli modelirat na način da bacamo sve 3 kocke odjednom

↳ tada je vjerojatnost:

$$N = 216$$

$$A = 1$$

$$P(A) = \frac{1}{216}$$

ZAD: Strijelac gađa metu 10 puta te je svaki put vjerojatnost pogotka 0.8 !

Odredi vjerojatnost da je strijelac:

a) svaki put promašio

b) točno 2 puta pogodio

c) barem 2 puta pogodio

d) najviše 2 puta pogodio

a) $P(A) = 0.2^{10}$

b) $P(B) = 0.8^2 \cdot 0.2^8 \cdot \binom{10}{2}$

↳ moramo odabrati 2 od 10 gađanja u kojima je strijelac pogodio

c) ⇒ razmišljamo suprotno:

$$P(C) = 1 - \binom{10}{1} \cdot 0.8 \cdot 0.2^9 - 0.2^{10}$$

d) $P(10) = 0.2^{10} + \binom{10}{1} \cdot 0.8 \; 0.2^9 + \binom{10}{2} \cdot 0.8^2 \cdot 0.2^8$

ZAD: U bubnju se nalaze 6 bijelih i 4 crne kuglice. Izvlačimo njih 7. Izračunaj da smo izvukli barem 5 bijelih kuglica ako:

a) izvlačimo odjednom 7 kuglica

b) izvlačimo jednu po jednu s vraćanjem

a) $$P(A) = \frac{\binom{6}{5}\cdot\binom{4}{2} + \binom{6}{6}\cdot\binom{4}{1}}{\binom{10}{7}}$$

b) ⇒ vidimo da je ovo sada ponavljanje pokusa jer je svako od izvlačenja neovisno o ostalima:

↳ primjeti da sada možemo izvući svih sedam bijelih kuglica

$$P(B) = \left(\frac{6}{10}\right)^5\left(\frac{4}{10}\right)^2\cdot\binom{7}{5} + \binom{7}{6}\left(\frac{6}{10}\right)^6\cdot\frac{4}{10} + \left(\frac{6}{10}\right)^7$$

BIRANJE POLOŽAJA BIJELIH KUGLICA

BESKONAČNI VJEROJATNOSNI PROSTOR

⇒ primjer ovakvog pokusa:

- bacamo kocku sve dok ne dobijemo šesticu
- strijelac gađa metu sve dok je ne pogodi

⇒ ALGEBRA DOGAĐAJA:

↳ algebra događaja $\mathcal{F}$ je σ-algebra ako vrijedi:

$$A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$$

⇒ tada vjerojatnost P zadovoljava svojstvo σ-aditivnosti:

$$\boxed{P\left(\bigcup_{n=1} A_n\right) = \sum_{n=1}^{\infty} P(A_n)}$$

, za disjunktne $A_1, A_2, \ldots$

⇒ Ω može biti:

a) prebrojiv

b) neprebrojiv (geometrijska vjerojatnost)

ZAD: Dva igrača, Ilko i Neven, naizmjence izvlače iz bubnja u kojem su dvije bijele i 6 crnih kuglica. Pobjednik je onaj koji prvi izvuče bijelu kuglicu. Izračunaj vjerojatnost pobjede oba igrača:

a) s vraćanjem u bubanj

b) bez vraćanja u bubanj

a) ⇒ modeliramo vjerojatnosni prostor:

$$\Omega = \{B, CB, CCB, CCCB, CCCCB, \dots\}$$

DOGAĐAJI ZA PRVOG IGRAČA

$$P(I) = \frac{2}{8} + \frac{6}{8}\cdot\frac{6}{8}\cdot\frac{2}{8} + \frac{6}{8}\cdot\frac{6}{8}\cdot\frac{6}{8}\cdot\frac{6}{8}\cdot\frac{2}{8} + \dots =$$

$$= \frac{1}{4} + \left(\frac{3}{4}\right)^2\cdot\frac{1}{4} + \left(\frac{3}{4}\right)^4\cdot\frac{1}{4} + \dots =$$

$$= \sum_{m=0}^{\infty} \frac{1}{4}\cdot\left(\frac{3}{4}\right)^{2m} = \frac{1}{4}\cdot\sum_{m=0}^{\infty}\left(\frac{9}{16}\right)^m$$

↳ geometrijski red: $\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m$, za $|x|<1$

$$P(I) = \frac{4}{7}$$

$$P(N) = 1 - P(I) \Rightarrow P(N) = \frac{3}{7}$$

b) ⇒ sada možemo naslutiti pravilo jer se stvari mijenjaju nakon prvog izvlačenja

$$P(I) = \frac{2}{8} + \frac{6}{8}\cdot\frac{5}{7}\cdot\frac{2}{6} + \frac{6}{8}\cdot\frac{5}{7}\cdot\frac{4}{6}\cdot\frac{3}{4}\cdot\frac{2}{4} + \frac{6}{8}\cdot\frac{5}{7}\cdot\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}$$

$$P(I) = \frac{4}{7}$$

$$P(N) = 1 - \frac{4}{7} \quad \Rightarrow \quad P(N) = \frac{3}{7}$$

↳ ovdje nam ispadne ista vjerojatnost, i to očekujemo da je tako u radacima gdje imamo 2 mogućnosti izvlačenja

ZAD: U bubnju se nalaze 14 kuglica s brojevima od 1 do 14. Izvlačimo 5 kuglica. Izračunaj vjerojatnost da je zbroj najvećeg dva broja (od izvučenih 5) barem 25 !

$$P(A) = \frac{\binom{10}{3} + \binom{11}{3} + \binom{11}{3} + \binom{12}{3}}{\binom{14}{5}}$$

ne smijemo izabrat broj već od 11, pa samo do 10 možemo

ovo su sve mogućnosti (redoslijed nije bitan)

$11 + 14 \geq 25$

$12 + 13 \geq 25$

$12 + 14 \geq 25$

$13 + 14 \geq 25$

ovo su povoljni događaji

ZAD: Problem rođendana! Izračunaj vjerojatnost da su barem dvije osobe od njih „n" rođene istog dana (datum u godini)!

⇒ razmišljamo suprotno:

$$P(A) = 1 - \frac{365 \cdot 364 \cdot 363 \cdot 362 \cdot \ldots \cdot (365 - n + 1)}{365^n}$$

↳ uvrstimo konkretne brojeve:

$n = 23 \longrightarrow 0,51$

$n = 40 \longrightarrow 0,9$

$n = 80 \longrightarrow 0,9999$

ZAD: Od 4 grla ra žarulje, 2 su ispravna, 2 su neispravna, a od 7 žarulja 4 su ispravne. Ako nasumično odaberemo 4 žarulje i nasumično ih stavimo u 4 grla, kolika je vjerojatnost da dobijemo svjetlo?

⇒ razmišljajmo suprotno:

↳ kada neće doći do svjetla

↳ to je zato jer je manje slučajeva kada neće biti svjetla - mora se dogoditi da na ispravna grla stavimo neispravne žarulje

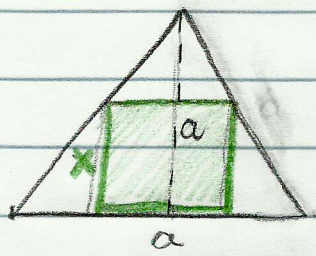
$$P(S) = 1 - \frac{\binom{3}{2} \cdot 2! \cdot \binom{5}{2} \cdot 2!}{\binom{7}{4} 4!}$$

neispravne žarulje na ispravna grla

bilo koje žarulje na neispravna grla

GEOMETRIJSKA VJEROJATNOST

ZAD: U jednakokračnom trokutu osnovica „a" i visina „a", upisan je kvadrat. Izračunaj vjerojatnost da slučajno odabrana točka u trokutu leži unutar kvadrata?



$$P(A) = \frac{P_{\square}}{P_{\triangle}} = \frac{x^2}{\frac{1}{2}a \cdot a} = \frac{\frac{a^2}{4}}{\frac{1}{2}} = \frac{1}{2}$$

sličnost trokuta

$$\frac{x}{a} = \frac{\frac{a}{2} - \frac{x}{2}}{\frac{a}{2}} \Rightarrow x = \frac{a}{2}$$

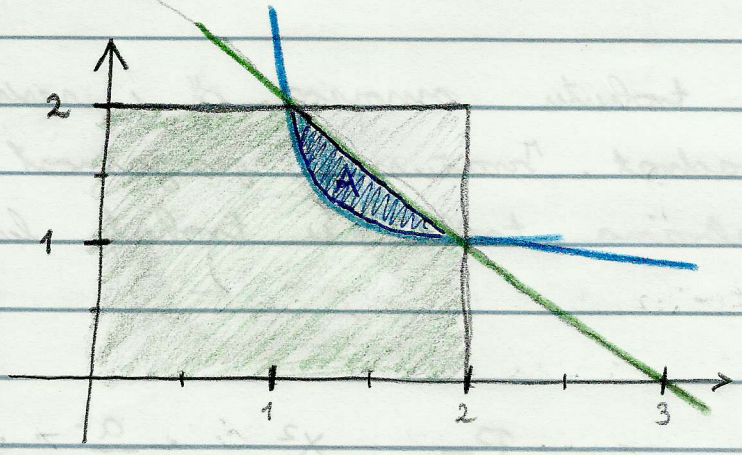
⇒ GEOMETRIJSKA VJEROJATNOST:

$$P(A) = \frac{m(A)}{m(\Omega)}$$; „m" → oznaka za mjeru

↳ jedna dimenzija → mjera je duljina
↳ dvije dimenzije → mjera je površina
↳ tri dimenzije → mjera je volumen

↳ lagano se pokaže da ova vjerojatnost zadovoljava svojstva vjerojatnosti

ZAD: Biramo dvije točke iz intervala $[0,2]$. Kolika je vjerojatnost da im je suma manja od 3, a produkt veći od 2?



$$y + x < 3$$

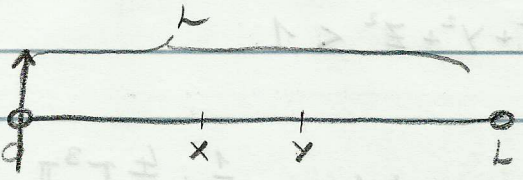
$$y < 3 - x$$

$$x \cdot y > 2 \Rightarrow y > \frac{2}{x}$$

$$P(A) = \frac{\int_1^2 \left(3 - x - \frac{2}{x}\right) dx}{4}$$

$$P(A) = \frac{\frac{3}{2} - 2\ln 2}{4}$$

ZAD: Štap duljine L je prelomljen na dva mjesta. Izračunajte vjerojatnost da je najkraći od njih manji od $\frac{1}{4}L$. Vjerojatnost preloma na svakom mjestu je jednaka!



↳ biramo dvije točke $x, y \in [0, L]$

↳ bitno je shvatiti da su:

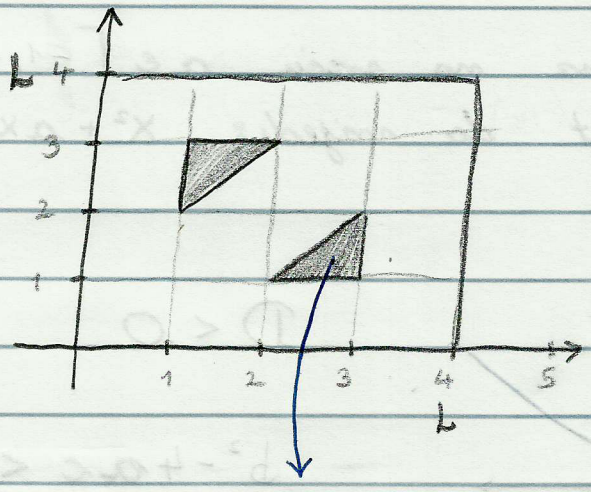
- ako je najmanji dio veći od $\frac{1}{4}L$, onda su sigurno i preostala dva dijela veća od $\frac{1}{4}L$

$\Rightarrow$ za $y > x$:

$x > \frac{1}{4}L$

$y - x > \frac{1}{4}L$

$L - y > \frac{1}{4}L$



$y > \frac{1}{4}L + x$

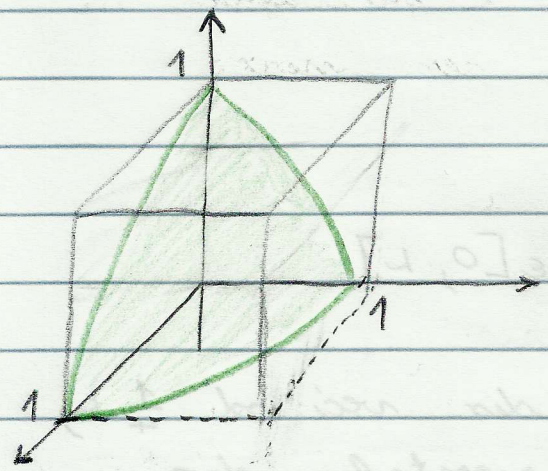
$y < \frac{3}{4}L$

$x > \frac{1}{4}L$

za slučaj $x > y$ (zbog simetrije problema)

$$P(A) = \frac{\left(\frac{1}{4}L\right)^2}{L^2} = \frac{1}{16}$$

ZAD: Biramo 3 točke iz intervala $[0,1]$. Izračunaj vjerojatnost da je zbroj njihovih kvadrata manji od 1 ?!

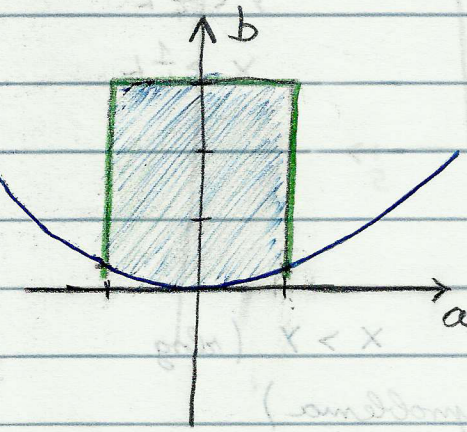


$$x^2 + y^2 + z^2 < 1$$

$$P(A) = \frac{V_A}{V_\Omega} = \frac{\frac{1}{8} \cdot \frac{4}{3} r^3 \pi}{1^3}$$

$$P(A) = \frac{\pi}{6}$$

ZAD: Traberima na sreću $a \in [-1,1]$, $b \in [0,3]$. Izračunaj vjerojatnost da vrijedi: $x^2 + ax + b > 0$, $\forall x \in \mathbb{R}$



$$D < 0$$

$$b^2 - 4ac < 0$$ } formula za diskriminantu

↳ uvjet za naš primjer: $a^2 - 4b < 0$

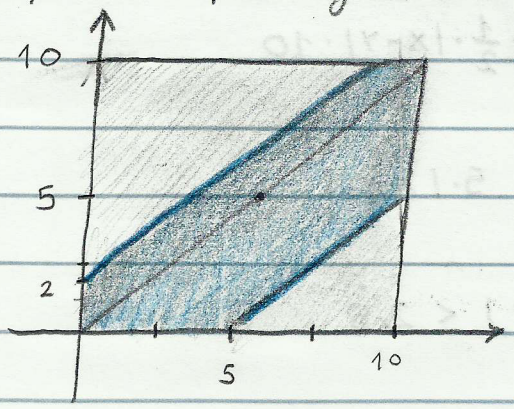
$$b > \frac{a^2}{4}$$

$$P(A) = \frac{6 - 2\int_0^1 \frac{x^2}{4} dx}{2 \cdot 3} \Rightarrow P(A) = \frac{6 - \frac{1}{6}}{6}$$

$$P(A) = \frac{35}{36}$$

ZAD: Lopov provaljuje u mjenjačnicu između 2:50 i 3 ujutro. Također, policajac obilazi tu mjenjačnicu između 2:50 i 3 ujutro. Lopovu za provalu treba 2 minute, a policajac se zadržava 5 minuta. Izračunaj vjerojatnost:

a) da policajac ulovi lopova





$X < Y < X + 2$

} slučaj da je prvo došao lopov

$Y > X$ $\quad$ $Y < X$

$Y < X + 2$ $\quad$ $Y > X - 5$



$Y < X < Y + 5$

} slučaj da je prvo došao policajac

$$P(A) = \frac{100 - \frac{1}{2} \cdot 5^2 - \frac{1}{2} \cdot 8^2}{100}$$

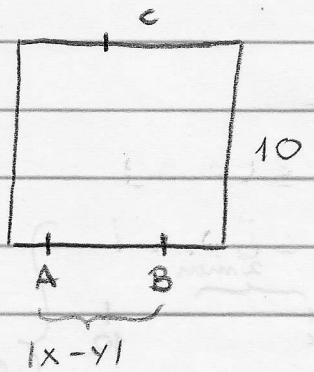
$$P(A) = \frac{111}{200}$$

b) izračunaj da lopov uspije provalit i otić prije nego što policajac dođe

$Y > X$
$X + 2 < Y$
} to je zapravo onaj gornji (veći) trokut

$$P(B) = \frac{\frac{1}{2} \cdot 8^2}{100} = 32\,\%$$

ZAD: Na jednoj strani kvadrata lijevo nasumice dvije točke A i B. Na suprotnoj strani lijevo točku C. Izračunaj vjerojatnost da je površina trokuta ABC manja od 25 ako je stranica kvadrata jednaka 10?

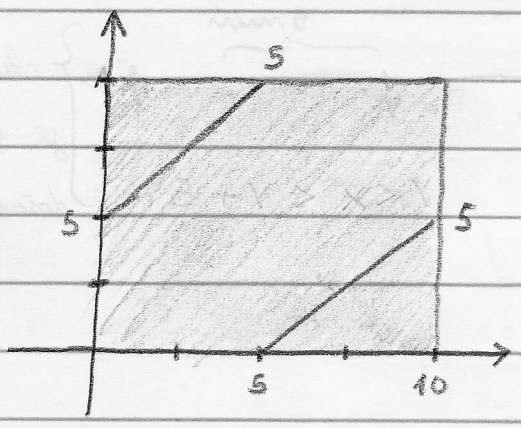


$$P_{\Delta ABC} = \frac{1}{2} \cdot |x-y| \cdot 10$$

$$P_{\Delta ABC} = 5 \cdot |x-y|$$

$$5 \cdot |x-y| < 25$$

$$|x-y| < 5$$



$y - x < 5$ $\quad$ $x - y < 5$

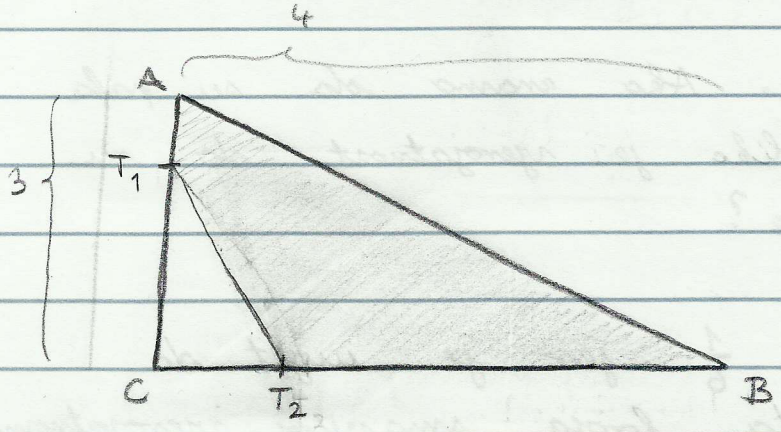
$y < x + 5$ $\quad$ $y > x - 5$

$$P(A) = \frac{100 - 5^2}{100}$$

$$P(A) = \frac{3}{4}$$

ZAD: U pravokutnom trokutu ABC duljine kateta su $|AC| = 3$, $|BC| = 4$. Točka T_1 se kreće na $\overline{AC}$, a T_2 na $\overline{BC}$. Odredi vjerojatnost da je zadovoljena relacija:

$$P_{\square ABT_1T_2} > \frac{1}{2} P_{\triangle ABC}$$



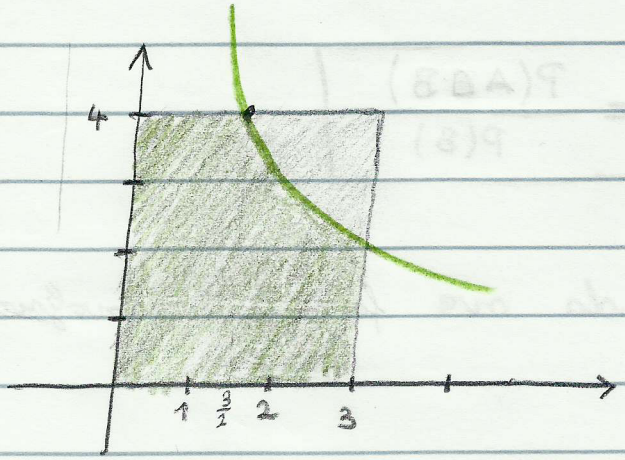
$x \in [0, 3]$

$y \in [0, 4]$

↳ početnu nejednadžbu zapisamo preko varijabli:

$$6 - \frac{1}{2} xy > \frac{1}{2} \cdot 6$$

$$xy < 6 \quad \Rightarrow \quad y < \frac{6}{x}$$



$$P(A) = \frac{4 \cdot \frac{3}{2} + \int_{3/2}^{3} \frac{6}{x}\, dx}{3 \cdot 4}$$

$$P(A) = \frac{1}{2} + \frac{1}{2} \ln 2$$

UVJETNA VJEROJATNOST

⇒ uvjetna vjerojatnost nam je potrebna kod problema kod kojih vjerojatnosni prostor nije „tako jednostavan"

PR: Bacili smo 2 kocke. Ako znamo da su pala 2 ista broja, kolika je vjerojatnost da su pale dvije šestice?

↳ vjerojatnost je $\frac{1}{6}$ jer je uvjet da su pala dva ista broja smanjio vjerojatnosni prostor na 6 mogućih ishoda

⇒ UVJETNA VJEROJATNOST:

↳ neka je B događaj takav da je $P(B) > 0$

↳ uvjetna vjerojatnost događaja A uz uvjet B se definira s:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

↳ lako se pokaže da ova formula zadovoljava uvjete vjerojatnosti

ZAD 5 U vrv se nalaze 4 plave, 5 bijelih i 6 crnih kuglica. Izvlačimo 3 kuglice jednu za drugom.

A = { sve kuglice su različite boje }

B = { prva kuglica je bijela }

Izračunaj vjerojatnost:

$$P(B|A) \quad i \quad P(A|B) \; ?$$

a)

$$P(B|A) = \frac{2}{3!} = \frac{1}{3}$$

BCP
BPC

b) $$P(A|B) = \frac{P(AB)}{P(B)}$$

$$P(A|B) = \frac{\frac{5}{15} \cdot \frac{4}{14} \cdot \frac{6}{13} + \frac{5}{15} \cdot \frac{6}{14} \cdot \frac{4}{13}}{\frac{5}{15} \cdot \frac{14}{14} \cdot \frac{13}{13}} = \frac{24}{91}$$

ZAD: Novčić bacamo 10 puta. Vjerojatnost da je svih 10 puta palo pismo, ako znamo da je u tih 10 bacanja pismo palo barem 9 puta?

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\left(\frac{1}{2}\right)^{10}}{\left(\frac{1}{2}\right)^{9} \cdot \frac{1}{2} \cdot \binom{10}{9} + \left(\frac{1}{2}\right)^{10}} = \frac{1}{11}$$

ZAD: 4 feroca i profesor izašli su van. Vjerojatnost uspješnog uleta od feroca je 0.3, a profesora 0.9. Ako znamo da je samo jedan let uspješan, izračunajte vjerojatnost da je to profesor?

$$P(A|B) = \frac{0.9 \cdot 0.7^4}{\binom{4}{1} \cdot 0.3 \cdot 0.7^3 \cdot 0.1 + 0.9 \cdot 0.7^4}$$

$$P(A|B) = 84\%$$

NEZAVISNOST

⇒ promotrimo sljedeći primjer koji je potreban kako bi ukazali da bitno formalnu definiciju nezavisnosti događaja

PR: U bubnju su 3 bijele i 7 crnih kuglica. Izvlačimo dvije kuglice jednu za drugom.

$A = \{$ prva izvučena je bijela $\}$

$B = \{$ druga izvučena je crna $\}$

Odredi nezavisnost događaja:

a) s vraćanjem ?

b) bez vraćanja ?

a) $P(A) = \frac{3}{10}$ $\qquad P(B) = \frac{7}{10}$

b) $P(A) = \frac{3}{10}$ $\qquad P(B) = \frac{3}{10} \cdot \frac{7}{9} + \frac{7}{10} \cdot \frac{6}{9}$

↳ primjeti da ako vraćamo kuglice, događaji A i B su međusobno nezavisni

↳ također, primjeti da vrijedi:

$$P(B \mid A) = P(B)$$

⇒ NEZAVISNOST DOGAĐAJA:

↳ za događaje A i B kažemo da su nezavisni ako vrijedi $P(A) = P(A|B)$ ili vrijedi $P(B) = P(B|A)$

⇒ definicija nezavisnosti ima neke korisne posljedice:

• nužan i dovoljan uvjet nezavisnosti:

$$P(A|B) = \frac{P(AB)}{P(B)} = P(A)$$

$$\Downarrow$$

$$P(AB) = P(A) \cdot P(B)$$

⇒ nezavisnost se najčešće pojavljuje kroz teoriju, a primjenjuje se u svim zadacima, ali prikrivena

ZAD: Ako su A i B nezavisni, dokaži da su $\bar{A}$ i $\bar{B}$ nezavisni?

$$P(\bar{A}\bar{B}) = P(\overline{A+B}) = 1 - P(A+B) =$$

$$= 1 - [P(A) + P(B) - P(AB)] =$$

$$= 1 - P(A) - P(B) + P(AB) =$$

$$= 1 - P(A) - P(B) + P(A) \cdot P(B) =$$

$$= 1 - P(A) + P(B)[P(A) - 1] =$$

$$= 1 - P(A) + P(B)[1 - P(A)] = [1 - P(A)] + [1 - P(B)] =$$

$$= P(\bar{A}) + P(\bar{B})$$

ZAD: Mogu li disjunktni događaji biti nezavisni?

$0 = P(AB) = P(A)P(B) \neq 0$ (za $P(A) > 0$, $P(B) > 0$)

↳ disjunktni događaji ne mogu biti nezavisni

ZAD: Ako je $P(B) \in \langle 0, 1 \rangle$ i $P(A|B) = P(A|\overline{B})$, dokažite da su A i B nezavisni!

$$P(A|B) = P(A|\overline{B})$$

$$\frac{P(AB)}{P(B)} = \frac{P(A\overline{B})}{P(\overline{B})}$$

$$\frac{P(AB)}{P(B)} = \frac{P(A) - P(AB)}{1 - P(B)}$$

$$P(AB) - P(B)P(AB) = P(A)P(B) - P(B)P(AB)$$

$$P(AB) = P(A)P(B)$$

⇒ pogledajmo sada općenitije, slučaj kada se pojavljuje više događaja:

⇒ NEZAVISNOST „M" DOGAĐAJA:

↳ događaji $A_1, A_2, \ldots, A_m$ su nezavisni ako za svaki izbor $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ za $2 \leq k \leq m$ vrijedi:

$$P(A_{i_1} \cdot \ldots \cdot A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$$

ZAD: Definiraj nezavisnost događaja A, B, C!

$$P(ABC) = P(A)\,P(B)\,P(C)$$

$$P(AB) = P(A)\,P(B)$$

$$P(AC) = P(A)\,P(C)$$

$$P(BC) = P(B)\,P(C)$$

} ako sve ove jednakosti vrijede, onda su A, B, C nezavisni događaji

ZAD: Ako su A, B, C nezavisni, dokažite da su A i $B \cup C$ nezavisni?

$$P(A(B \cup C)) = P(A)\,P(B \cup C)$$

ovo trebamo dobit

$$P(A \cdot (B \cup C)) = P(AB \cup AC) = P(AB) + P(AC) - P(ABC) =$$

$$= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) =$$

$$= P(A) \cdot \left[P(B) + P(C) - P(B)P(C) \right] =$$

$$= P(A) \cdot P(B \cup C)$$

POTPUNA VJEROJATNOST

PR: Student odgovara na pitanje s 4 ponuđena odgovora od kojih je jedan točan. Vjerojatnost da zna odgovor je 80%, a ako ne zna, student raskrižuje na sreću (nema negativnih bodova). Izračunaj vjerojatnost da je student točno odgovorio?

↳ H_1, H_2 su hipoteze

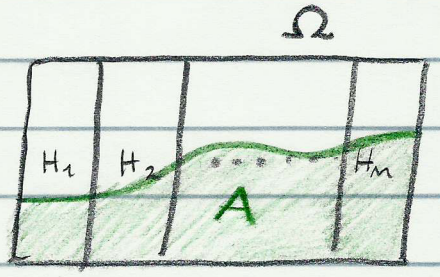
$$P(A) = \underbrace{0.8}_{P(H_1)} + \underbrace{\frac{1}{4} \cdot 0.2}_{P(H_2)} = 0.85\,\%$$

Ako je točno odgovorio, kolika je vjerojatnost da student nije pogađao odgovor?

$$P(H_1 \mid A) = \frac{P(H_1 A)}{P(A)} = \frac{0.8}{0.85} = 0.941$$

⇒ POTPUNA VJEROJATNOST:

↳ vjerojatnosni prostor potpuno rascjepamo na hipoteze



↳ primjetimo da su hipoteze disjunktne

↳ *izračunajmo sada* $P(A)$:

$$P(A) = P(AH_1 \cup AH_2 \cup \dots \cup AH_m) =$$

$$= P(AH_1) + P(AH_2) + \dots + P(AH_m) =$$

$$= P(H_1)\,P(A|H_1) + P(H_2)\,P(A|H_2) + \dots + P(H_m)\,P(A|H_m)$$

↳ *općenita formula*:

$$\boxed{P(A) = \sum_{i=1}^{n} P(H_i)\,P(A|H_i)}$$

↳ *uvjet da vrijedi*:

$$P(H_1) + P(H_2) + \dots + P(H_m) = 1$$

ZAD: Kod testiranja na određenu bolest u populaciji, ako ispitanik ima bolest test je pozitivan u 90% slučajeva, a ako nema, test je pozitivan u 5% slučajeva. U prosjeku, 2% populacije ima tu bolest. Kolika je vjerojatnost da će odabrana osoba iz populacije test biti pozitivan?

$$P(A) = 0.02 \cdot 0.9 + 0.98 \cdot 0.05$$

$$P(A) = 0.067$$

$A \rightarrow$ test pozitivan

$H_0 \rightarrow$ nema bolest

$H_1 \rightarrow$ ima bolest

$$P(A) = P(H_1) \cdot P(A|H_1) + P(H_0) \cdot P(A|H_1)$$

Ako je test bio pozitivan, kolika je vjerojatnost da je osoba bolesna?

$$P(H_1|A) = \frac{P(A \cap H_1)}{P(A)} = \frac{0.02 \cdot 0.9}{0.067}$$

$$P(H_1|A) = 0.2687$$

⇒ BAYESOVA FORMULA:

$$P(H_i \mid A) = \frac{P(H_i)\,P(A \mid H_i)}{P(A)}$$

ovo je samo preuređena formula uvjetne vjerojatnosti

ZAD: Izvor emitira znak 1 sa vjerojatnošću 0.75, a 0 sa 0.25. Na izlazu, 5% znakova se pogrešno interpretira. Kolika je vjerojatnost da smo na izlazu primili jedinicu?

$$P(A) = 0.75 \cdot 0.95 + 0.25 \cdot 0.05$$

$$P(A) = 0.725$$

Ako je primljena jedinica, kolika je vjerojatnost da je ona i poslana?

$$P(H_1 \mid A) = \frac{P(H_1)\,P(A \mid H_1)}{P(A)}$$

$$P(H_1 \mid A) = 0.983$$

ZAD: U kutiji imamo 2 ispravne i 2 lažne kocke na kojoj su sve brojeve šestice. Nasumično izvlačimo dvije kocke i bacimo ih. Ako su pale dvije šestice, odredi vjerojatnost da smo izvukli dvije lažne kocke?

↳ imamo 2 slučajna pokusa
- ishod drugog ovisi o ishodu prvog pokusa ⇒ odmah razmišljaj o potpunoj vjerojatnosti

$H_0 = \{2$ ispravne (I), 0 lažnih $(L)\}$
$H_1 = \{1\ I\ , 1\ L\}$
$H_2 = \{2\ I\ , 0\ L$
$A = \{$ pale su dvije $6\}$

$P(H_2 \mid A) = ?$

$$P(H_0) = \frac{\binom{2}{2}}{\binom{4}{2}} = \frac{1}{6} \quad \Rightarrow \quad P(A \mid H_0) = \frac{1}{6} \cdot \frac{1}{6}$$

$$P(H_1) = \frac{\binom{2}{1}\binom{2}{1}}{\binom{4}{2}} = \frac{4}{6} \quad \Rightarrow \quad P(A \mid H_1) = \frac{1}{6} \cdot 1$$

$$P(H_2) = \frac{1}{6} \quad \Rightarrow \quad P(A \mid H_2) = 1$$

$$P(H_2 \mid A) = \frac{P(H_2)\,P(A \mid H_2)}{P(A)} = \frac{P(H_2)\,P(A \mid H_2)}{\sum P(H_i)\,P(A \mid H_i)} = \frac{36}{61}$$

ZAD: U gradu ima 1000 ljudi od kojih 950 uvijek govori istinu, a 50 uvijek govori laži. Policija ima detektor koji griješi u 1% slučajeva. Ako je detektor utvrdio da on laže, koja je vjerojatnost da je on lažac?

$H_0 = \{$ uhvaćen lažac $\}$

$H_1 = \{$ uhvaćen iskren čovjek $\}$

$P(H_0) = \frac{1}{20} = 0.05 \qquad P(H_1) = 0.95$

$P(A) = \{$ detektor pokaraca lažl $\}$

$$P(H_0 \mid A) = \frac{P(H_0 A)}{P(A)} = \frac{P(A) \cdot P(A \mid H_0)}{\sum P(H_i \mid P(A \mid H_i)}$$

$P(H_0 \mid A) = 0.839$

ZAD: Krvne grupe: 34 % ima grupu O
37 % ima grupu A
21 % ima grupu B
8 % ima grupu AB

a) Kolika je vjerojatnost da slučajno odabrana osoba može primiti krv od druge slučajne?

H_O → izabral smo O $\quad P(H_O) = 0.34$

H_A → izabral smo A $\quad P(H_A) = 0.37$

H_B → izabral smo B $\quad P(H_B) = 0.21$

H_{AB} → izabral smo AB $\quad P(H_{AB}) = 0.08$

$U = \{$ uspjetna transfurzija $\}$

$P(U|H_O) = 0.34$

$P(U|H_A) = 0.34 + 0.37$

$P(U|H_B) = 0.21 + 0.34$

$P(U|H_{AB}) = 1$

$P(U) = \sum P(H_i) P(U|H_i)$

$P(U) = 0.5738$

b) Ako je transfurzija uspjela, koja je vjerojatnost da je bila grupa AB?

$$P(H_{AB}|U) = \frac{P(H_{AB}) P(U|H_{AB})}{P(U)} = 0.1394$$

ZAD: 4 lovca gadaju u metu. Vjerojatnost pogotka su 0.3, 0.2, 0.2, 0.2. Da bi ubili vepra, potrebna su 2 pogotka. Jedan od lovaca je sparao vepra, gadao 3 puta i ubio vepra. Izračunaj vjerojatnost da je lovac 1 gadao?

$H_1 = \{ \text{prvi lovac} \}$ $\quad P(H_i) = \frac{1}{4}$

$\vdots$

$H_i = \{ \text{"}i\text{"-ti lovac} \}$ $\quad A = \{ \text{vepar je ubijen} \}$

$$P(A \mid H_1) = 0.3^2 \cdot 0.7 \cdot \binom{3}{2} + 0.3^3$$
$$P(A \mid H_i) = 0.2^2 \cdot 0.8 \cdot \binom{3}{2} + 0.2^3, \quad i = 2, 3, 4$$

$$P(H_1 \mid A) = \frac{P(H_1)\, P(A \mid H_1)}{\sum P(H_i)\, P(A \mid H_i)} = 0.41$$

ZAD: U tri kutije nalaze se crne i bijele kuglice. U prvoj kutiji je 3 puta više bijelih nego crnih, a u ostale dvije je 3 puta više crnih nego bijelih. Nasumično izaberemo jednu kutiju, i iz nje izvlačimo 5 kuglica jednu po jednu s vraćanjem. Ako smo izvukli 3 bijele i 2 crne, kolika je vjerojatnost da smo izabrali iz prve kutije?

$H_i = \{$ odabrali smo „i"-tu kutiju$\}$

$A = \{$ izvukli smo 3B i 2C$\}$

$$P(H_i) = \frac{1}{3}$$

$$P(H_1 \mid A) = ?$$

$$P(H_1 \mid A) = \frac{P(H_1 A)}{P(A)} = \frac{P(H_1)\, P(A \mid H_1)}{P(A)}$$

$$P(H_1 \mid A) = \frac{P(H_1)\, P(A \mid H_1)}{\sum P(H_i)\, P(A \mid H_i)}$$

$$P(A \mid H_1) = \binom{5}{3} \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right)^2$$

$$P(A \mid H_i) = \binom{5}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 \qquad i = 2, 3$$

DISKRETNE SLUČAJNE VARIJABLE

⇒ neka je $S = \{x_1, x_2, \ldots\}$ konačan ili prebrojiv skup, te neka varijabla X svakom elementarnom događaju nekog slučajnog pokusa pridružuje vrijednost iz S

↳ ravnina nas vjerojatnost da X poprimi neku vrijednost x_k, tj. ravnina nas vjerojatnost skupa $A_k = \{\omega \in \Omega : X(\omega) = x_k\}$

↳ preslikavanje $X : \Omega \to S$ je DISKRETNA SLUČAJNA VARIJABLA ako $\forall x_k \in S$ je skup $A_k \in \mathcal{F}$ (tj. ako je A_k događaj!)

⇒ OZNAKE:

$$p_k = P(X = x_k), \quad p_k > 0 \quad \sum p_k = 1$$

↳ zakon razdiobe pišemo ovako:

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$$

PR: Neka je X broj na kocki $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$

↳ primjet da je ovo jednolika razdioba

$$P(X \geq 4) = P(X=4) + P(X=5) + P(X=6) = \frac{1}{2}$$

$$P(X > 2) = 1 - \frac{2}{6} = \frac{4}{6} = \frac{2}{3}$$

ZAD: Bacamo 3 kocke. Neka je X broj šestica. Odredi zakon razdiobe od X!

$$X \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \left(\frac{5}{6}\right)^3 & 3 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 & \binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} & \left(\frac{1}{6}\right)^3 \end{pmatrix}$$

⇒ SVOJSTVA SLUČAJNIH VARIJABLI:

↳ slučajne varijable $X, Y : \Omega \to S$ su nezavisne za sve $x_k, y_j \in S$, ako vrijedi:

$$P(X = x_k, Y = y_j) = P(X = x_k) \cdot P(Y = y_k)$$

↳ **općenito**: slučajne varijable $X_1, X_2, \ldots, X_m$ na istom Ω su nezavisne ako za sve $A_1, A_2, \ldots, A_m \subset S$ vrijedi:

$$\boxed{P(X_1 \in A_1, \ldots, X_m \in A_m) = P(X_1 \in A_1) \cdot \ldots \cdot P(X_m \in A_m)}$$

⇒ funkcije diskretne slučajne varijable

↳ neka je X diskretna slučajna varijabla

i neka je $\Psi : R \to R$ neka funkcija

↳ varijabla od $Y = \Psi(X)$ glasi:

$$Y \sim \begin{pmatrix} \Psi(x_1) & \Psi(x_2) & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

DISKRETNI SLUČAJNI VEKTORI

⇒ mi se bavimo samo dvodimenzionalnim slučajnim vektorima

⇒ razdioba slučajnog vektora (X, Y) je definirana s

$$p_{ij} = P(X = x_i, Y = y_j)$$

↳ ove razdiobe zapisujemo ovako:

$X \backslash Y$	y_1	y_2	...	y_j	...	y_m	
x_1	p_{11}	p_{12}				p_{1m}	p_1
x_2	p_{21}	p_{22}				p_{2m}	p_2
⋮							⋮
x_i	----------			p_{ij}		------	p_i
⋮							⋮
x_n	p_{n1}	p_{n2}	.....			p_{nm}	p_n
	q_1	q_2	...	q_j	...	q_m	

$$\sum \sum p_{ij} = 1$$

⇒ X i Y su nezavisni ako $p_{ij} = P(X = x_i) \cdot P(Y = y_j)$

↳ pomoću gornjeg zapisa dobivamo:

$$\boxed{p_{ij} = p_i \cdot q_j} \quad \forall i, j$$

↳ „p" i „q" su **marginalne razdiobe**

↳ to je kao da promatramo razdiobu samu za sebe

ZAD: Bacimo dvije kocke. Neka je X broj na prvoj kocki, a Y veći od dvaju brojeva koje smo dobili.

a) Napiši zakon razdiobe vektora (X, Y)

X \ Y	1	2	3	4	5	6	
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
2	0	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
3	0	0	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
4	0	0	0	$\frac{4}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
5	0	0	0	0	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
6	0	0	0	0	0	$\frac{6}{36}$	$\frac{6}{36}$
	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	1 ✓ OK

b) Odredi marginalnu razdiobu od X!

↳ marginalna razdioba od X:

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{6}{36} & \frac{6}{36} & \frac{6}{36} & \frac{6}{36} & \frac{6}{36} & \frac{6}{36} \end{pmatrix}$$

ZAD: Zadana je razdioba vektora (X, Y)!

$X \backslash Y$	0	1	
-1	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{10}{24}$
0	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{7}{24}$
1	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{7}{24}$
	$\frac{13}{24}$	$\frac{11}{24}$	1

a) Odredi marginalne razdiobe!

b) Ispitaj nezavisnost!

↳ vidimo da su varijable zavisne jer:

$$p_{11} \neq p_1 \cdot q_1$$

c) Odredi vjerojatnost $P(X \geq 0 \mid Y = 1) = ?$

$$P(X \geq 0 \mid Y > 1) = \frac{P(X \geq 0, Y = 1)}{P(Y = 1)} = \frac{\frac{7}{24}}{\frac{11}{24}} = \frac{7}{11}$$

d) Odredi razdiobu od (Z, W) ako je $Z = X + Y$, a $W = X \cdot Y$!

$Z \backslash W$	-1	0	1
-1	0	$\frac{1}{4}$	0
0	$\frac{1}{6}$	$\frac{1}{6}$	0
1	0	$\frac{1}{4}$	0
2	0	0	$\frac{1}{6}$

$Z = -1$:

$X = -1, Y = 0 \Rightarrow W = 0$

$Z = 2$:

$X = 1, Y = 1 \Rightarrow W = 1$

$Z = 1$:

$X = 1, Y = 0 \Rightarrow W = 0$

$X = 0, Y = 1 \Rightarrow W = 0$

$Z = 0$

$X = 0, Y = 1 \Rightarrow W = 0$

$X = 1, Y = 0 \Rightarrow W = 0$

ZAD: Neka je $X \sim \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, $Y \sim \begin{pmatrix} 1 & 2 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$

X, Y su nezavisne slučajne varijable. Odredi zakon vektora (X, Y), zakon varijable $Z = \frac{X}{Y}$?

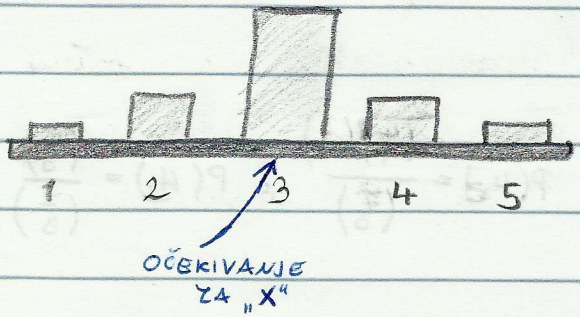
$X \backslash Y$	1	2	
1	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
2	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
	$\frac{1}{3}$	$\frac{2}{3}$	

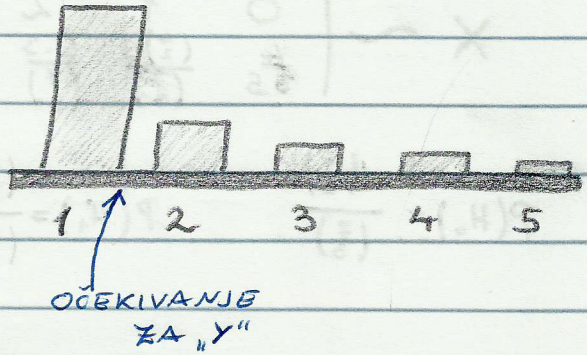
} ove sume izračunate su marginalne jer su varijable nezavisne

$$Z \sim \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ \underbrace{\frac{1}{6}+\frac{1}{3}}_{1=\frac{1}{1}\text{ ili }\frac{2}{2}} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

KARAKTERISTIKE SLUČAJNIH VARIJABLI

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0{,}05 & 0{,}2 & 0{,}5 & 0{,}2 & 0{,}05 \end{pmatrix} \qquad Y \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0{,}7 & 0{,}2 & 0{,}05 & 0{,}03 & 0{,}02 \end{pmatrix}$$





⇒ OČEKIVANJE diskretne slučajne varijable definira se kao:

$$E(X) = \sum_i x_i p_i$$

(ako taj red konvergira)

⇒ očekivanje stoga nije nužno cijeli broj te se ne zaokružuje (u većini zadataka)

ZAD: U kutiji se nalaze 3 crvene i 4 zelene kuglice. Izvlačimo 3 kuglice. Odredi očekivan broj crvenih kuglica ∇

$X \Rightarrow$ broj crvenih kuglica

$$X \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{4}{35} & \frac{\binom{4}{2}\cdot 3}{\binom{7}{3}} & \frac{4\cdot 3}{\binom{7}{3}} & \frac{1}{\binom{7}{3}} \end{pmatrix}$$

$$P(H_0) = \frac{\binom{4}{3}}{\binom{7}{3}} \qquad P(H_1) = \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} \qquad P(H_2) = \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}} \qquad P(H_3) = \frac{\binom{3}{3}}{\binom{7}{3}}$$

$$E(X) = \frac{9}{7}$$

Zad: Imamo 10 numeriranih kuglica (od nula do deset). Izvlačimo 4 kuglice. Odredi očekivanje od:

$X \Rightarrow$ drugi najveći izvučeni broj

$$X \sim \begin{pmatrix} 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \frac{7}{\binom{10}{4}} & \frac{\binom{3}{2}\binom{6}{1}}{\binom{10}{4}} & \frac{\binom{4}{2}\cdot 5}{\binom{10}{4}} & \frac{\binom{5}{2}\cdot 4}{\binom{10}{4}} & \frac{\binom{6}{2}3}{\binom{10}{4}} & \frac{\binom{7}{2}\cdot 2}{\binom{10}{4}} & \frac{1\cdot\binom{8}{2}}{\binom{10}{4}} \end{pmatrix}$$

$$P(X=k) = \frac{(10-k)\binom{k-1}{2}}{\binom{10}{4}}$$

$$E(X) = 6.6$$

ZAD: U kutiji se nalaze 1 bijela i 4 crne kuglice. Izvlačimo jednu po jednu dok ne izvučemo bijelu. Neka je X broj izvlačenja. Odredi očekivani broj izvlačenja:

a) bez vraćanja

b) s vraćanjem

a)

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{1}{5} & \frac{4}{5}\cdot\frac{1}{4} & \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{4} & \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2} & \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 \end{pmatrix}$$

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix} \Rightarrow E(X) = 3$$

b)

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \dots & m & \dots \\ \frac{1}{5} & \frac{4}{5}\cdot\frac{1}{5} & \left(\frac{4}{5}\right)^2\cdot\frac{1}{5} & \left(\frac{4}{5}\right)^3\cdot\frac{1}{5} & \left(\frac{4}{5}\right)^4\cdot\frac{1}{5} & \dots & \left(\frac{4}{5}\right)^{m-1}\cdot\frac{1}{5} & \dots \end{pmatrix}$$

$$E(X) = \sum_{m=1}^{\infty} m\cdot\left(\frac{4}{5}\right)^{m-1}\cdot\frac{1}{5} = \frac{1}{5}\sum_{m=1}^{\infty} m\left(\frac{4}{5}\right)^{m-1}$$

$$\sum x^m = \frac{1}{1-x} \Big/ \quad \Rightarrow \sum m x^{m-1} = \frac{1}{(1-x)^2}$$

$$E(X) = \frac{1}{5}\cdot\frac{1}{\left(1-\frac{4}{5}\right)^2} = 5 \;//$$

⇒ SVOJSTVA OČEKIVANJA:

↳ neka su X i Y slučajne varijable na istom vjerojatnosnom prostoru:

1) $\forall s, t \in \mathbb{R}$ vrijedi $E(sX + tY) = sE(X) + tE(Y)$

2) ako su X i Y nezavisne, tada vrijedi:

$$E(XY) = E(X) \cdot E(Y)$$

• **DOKAZ 1.**

↳ $E(sX) = \sum_i sx_i p_i = s\sum_i x_i p_i = s \cdot E(X)$ //

↳ $E(X+Y) = \sum_i \sum_j (x_i + y_j) p_{ij} =$

$$= \sum_i \underbrace{\sum_j x_i p_{ij}}_{p_i} + \sum_i \sum_j y_j p_{ij} =$$

$$= \sum_i x_i p_i + \sum_j y_j p_j = E(X) + E(Y)$$ //

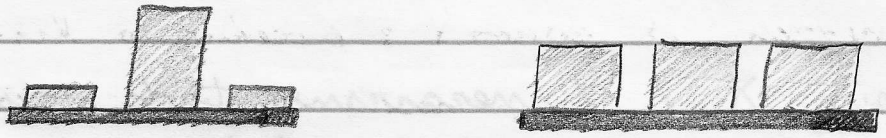
• **DOKAZ 2.**

↳ $E(XY) = \sum_i \sum_j x_i y_j p_{ij} = \sum_i \sum_j x_i y_j p_i q_j =$

$$= \sum_i x_i p_i \cdot \sum_j y_j q_j = E(X) \cdot E(Y)$$ //

⇒ DISPERZIJA (UVOD):

↳ moguće je da raspodjele imaju ista očekivanja, a u svitku su skroz različite:



↳ stoga, trebamo još jednu „veličinu" koja nam opisuje odstupanje razdiobe od očekivanja

⇒ neka je poznata razdioba X i neka je $Y = \Psi(X)$

↳ tada je očekivanje od Y:

$$E(Y) = E(\Psi(X)) = \sum_i \Psi(x_i)\, p_i$$

↳ najvažniji slučajevi:

a) $Y = X^m$ b) $Y = (X - \underbrace{E(X)}_{\text{neki broj}})^m$

$$E(X^m) = \sum x_i^m p_i \rightarrow \text{ishodišni moment reda „}m\text{"}$$

$$E\left((X - E(X))^m\right) = \sum (x_i - E(X))^m p_i \rightarrow \text{centralni moment reda „}m\text{"}$$

⇒ DISPERZIJA slučajne varijable X je centralni moment reda 2:

$$D(X) = E\left((X - E(X))^2\right)$$

↳ disperzija je mjera: „očekivana kvadratna udaljenost od očekivanja"

↳ UOČIMO: $D(X) \geq 0$

⇒ STANDARDNA DEVIJACIJA:

$$\boxed{\sigma = \sqrt{D(X)}}$$

PR: Dokažimo da vrijedi: $D(X) = E(X^2) - E(X)^2$

↳ ova je formula koja se najčešće koristi u radacima s praksi

↳ označimo: $\bar{X} = E(X)$

$$D(X) = E\left((X - \bar{X})^2\right) = E(X^2 - 2X\bar{X} + \bar{X}^2) =$$

$$= E(X^2) - 2\bar{X}\underbrace{E(X)}_{\bar{X}} + \bar{X}^2 = E(X^2) - \bar{X}^2 \quad //$$

⇒ SVOJSTVA DISPERZIJE:

1) $D(\lambda X) = \lambda^2 D(X)$

2) za nezavisne X i Y vrijedi: $D(X+Y) = D(X) + D(Y)$

• **DOKAZ 1**

↳ $D(\lambda X) = E\left((\lambda X)^2\right) - \left(E(\lambda X)\right)^2 = \lambda^2 E(X^2) - \lambda^2 E(X)^2$

$= \lambda^2 \left(E(X^2) - E(X)^2 \right) = \lambda^2 D(X)$

• **DOKAZ 2**

↳ $D(X+Y) = E\left((X+Y)^2\right) - \left(E(X+Y)\right)^2 =$

$= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 =$

za nezavisne X, Y: $E(XY) = E(X)E(Y)$

$= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 =$

$= D(X) + D(Y)$

ZAD: Zadana je slučajna varijabla:

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & k & \dots \\ p & p^2 & p^3 & p^4 & \dots & p^k & \dots \end{pmatrix}$$

Odredite:

p, $E(X)$, $D(X)$!

$$\sum_{k=1}^{\infty} p^k = \frac{p}{1-p} = 1 \Rightarrow p = \frac{1}{2}$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot p^k = 2$$

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \Big/ \; '$$

→ derivirajem se gubi prvi član (sada „n" kreće od 1)

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2} \quad \Big/ \cdot x$$

$$\sum_{n=1}^{\infty} n x^n = \frac{x}{(1-x)^2} \longrightarrow \sum_{n=1}^{\infty} n^2 x^{n-1} = \frac{(1-x)^{\cancel{2}} + 2x\cancel{(1-x)}}{(1-x)^{\cancel{4}\,3}} \quad \Big/ \cdot x$$

$$\sum_{n=1}^{\infty} n^2 x^n = x \cdot \frac{1-x+2x}{(1-x)^3}$$

$$D(X) = \sum_{k=1}^{\infty} k^2 p^k - E(X)^2$$

$$D(X) = 6 - 4 = 2$$

MEĐUSOBNI ODNOS DVIJE VARIJABLE

ZAD: Od 500 studenata koj su slušal ViS i MAT3, 370 ih je položilo oboje, 30 ih je položilo samo ViS, a 10 ih je položilo samo MAT3 i palo ViS! Postoj l neka vera (korelacija) izmedu polaganja ViS-a i MAT3?

↳ prvo uzjek činimo slučajni vektor:

$X = \{ViS\}$ $Y = \{MAT3\}$

0 = pad 1 = prolaz

X\Y	0	1	
0	0.18	0.02	0.2
1	0.06	0.74	0.8
	0.24	0.76	

$E(X) = 0.8$ $D(X) = 0.16$

$E(Y) = 0.76$ $D(Y) = 0.1824$

$$E(XY) = \sum\sum x_i y_j \cdot p_{ij} = 0.74$$

$$cov(X,Y) = 0.132$$

$$\tau(X,Y) = \frac{0.132}{\sqrt{0.16 \cdot 0.1824}} = 0.77$$

↑ polaganje MA3 i Vi S-a je dobro korelirano

⇒ KOVARIJACIJSKI MOMENT varijable X i Y definira se formulom:

$$cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

↳ izmnožimo gornji izraz i primijenimo linearnost očekivanja i dobijemo formulu koju ćemo puno više koristiti:

$$\boxed{cov(X,Y) = E(XY) - E(X)E(Y)}$$

↳ UOČI: ako su X i Y nezavisne, tada je $cov(X,Y) = 0$ (OBRAT NE VRIJEDI!?)

⇒ u praksi se puno češće koristi KOEFICIJENT KORELACIJE koji se definira ovako:

$$\boxed{\tau(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}}$$

ZAD: Neka je X jednak zbroju na kocki, a Y je jednak 0 ako je broj paran, a 1 ako je broj neparan. Bacamo kocku. Odredi korelaciju X i Y.

X \ Y	0	1	
2	0	$\frac{1}{6}$	$\frac{1}{6}$
4	$\frac{1}{6}$	0	$\frac{1}{6}$
6	0	$\frac{1}{6}$	$\frac{1}{6}$
8	$\frac{1}{6}$	0	$\frac{1}{6}$
10	0	$\frac{1}{6}$	$\frac{1}{6}$
12	$\frac{1}{6}$	0	$\frac{1}{6}$
	$\frac{1}{2}$	$\frac{1}{2}$	

$E(X) = 7$

$E(Y) = \frac{1}{2}$

$E(XY) = 3$

↳ sada možemo dobit kovarijacijski moment:

$$cov(X, Y) = -\frac{1}{2}$$

$$D(X) = \sum x^2 p_i - 7^2 = \frac{35}{3} \qquad D(Y) = \frac{1}{4}$$

↳ sada možemo izračunat koeficijent korelacij:

$$\tau(X, Y) = \frac{-\frac{1}{2}}{\sqrt{\frac{35}{3} \cdot \frac{1}{4}}}$$

$$\tau(X, Y) = -0.293$$

⇒ TEOREM O DISPERZIJI ZBROJA:

$$D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D(X_i) + 2\sum_{i<j} cov(X_i, Y_j)$$

↳ Dokaz:

$$D\left(\sum_{i=1}^{n} X_i\right) = E\left(\sum_{i=1}^{n} X_i - \underbrace{E\left(\sum_{i=1}^{n} X_i\right)}_{\sum_{i=1}^{n} E(X_i)}\right)^2 =$$

; $\sum X_i = \overline{X_i}$

$$= E\left(\sum_{i=1}^{\infty} (X_i - \overline{X_i})\right)^2 = \sum_{i=1}^{\infty} E(X_i - \overline{X_i})^2 + 2\sum_{i<j}^{n} E(X_i - \overline{X_i})(X_j - \overline{X_j}) =$$

$$= \sum_{i=1}^{\infty} D(X_i) + 2\sum_{i<j}^{n} cov(X_i, Y_j)$$ ∎

↳ znamo već da ako su $X_1, \ldots, X_n$ nezavisne, tada je $cov(X_1, \ldots, X_n) = 0$, tj.:

$$D\left(\sum_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} D(X_i)$$

⇒ CENTRIRANJE I NORMIRANJE SLUČAJNE VARIJABLE:

↳ ako znamo varijablu od X, tada znamo i varijablu od „X-a", $a \in R$:

$$E(X-a) = E(X) - a \quad , \quad D(X-a) = D(X)$$

↳ <u>bitan zaključak je da se ne mijenja mjt:</u>

$$cov(X-a, Y-b) = E\big(x-a-E(x-a)\big)\big(Y-b-E(Y-b)\big) =$$

$$= E\big(X-E(x)\big)\big(Y-E(Y)\big) =$$

$$= cov(X,Y) \;//$$

↳ <u>ne mijenja se mjt:</u>

$$\tau(X-a, Y-b) \;//$$

↳ pomoću svojstva „ne-promjene" zbog pomicanja, mi želimo centrirati varijablu (pomaknut ju u ishodište):

↳ odaberemo „a" takav da:

$a = E(x) = \bar{X}$ jer je tada:

$$E(X-\bar{X}) = 0 \quad ; \quad D(X-\bar{X}) = D(X)$$

↳ centrirana slučajna varijabla:

$$\dot{X} = X - \bar{X}$$

⇒ O ŽELJENOJ DISPERZIJI:

↳ promatramo varijablu: $aX + b$

↳ tada je:

$$E(aX+b) = aE(X) + b$$

$$D(aX+b) = a^2 D(X)$$

↳ odavde slijedi:

$$\boxed{X^* = \frac{X - \bar{x}}{\sigma_x}}$$ → tada je X normirana i centrirana varijabla (sve podatke u statistici su normiran i centriran)

↳ dokaz:

$$E(X^*) = \frac{1}{\sigma} E(X - \bar{x}) = 0$$

$$D(X^*) = \frac{1}{\sigma^2} D(X - \bar{x}) = 1$$

⇒ UOČIMO: normiranje ne mijenja koeficijent korelacije!

↳ dokaz:

$$r(X^*, Y^*) = \frac{E(X^*Y^*) - E(X^*)E(Y^*)}{\sigma(X^*)\sigma(Y^*)} = E\left(\frac{X-\bar{x}}{\sigma_x} \cdot \frac{Y-\bar{y}}{\sigma_y}\right) =$$

($\sigma(X^*)\sigma(Y^*) = 1$)

$$= \frac{cov(X,Y)}{\sigma_x \sigma_y} = r(X,Y)$$

⇒ TEOREM:

a) dokažimo slijedeće: $|r(X,Y)| \leq 1$

$$D(X^* \pm Y^*) = D(X^*) + D(Y^*) \pm 2\,\mathrm{cov}(X^*, Y^*)$$

$$= 2\left[1 \pm r(X,Y)\right]$$

teška dosjetka!

↳ za disperziju znamo da je uvijek veća il jednaka od nule:

$$2\left[1 \pm r(X,Y)\right] \geq 0$$

$$|r(X,Y)| \leq 1$$

b) dokažimo da vrijedi $|r(X,Y)| = 1$, ako je $Y = aX + b$ ∇

↳ vidimo da jednakost vrijedi ako je

$$D(X^* + Y^*) = 0$$

↳ disperzija je nula samo ako je varijabla konstanta:

$$X^* + Y^* = \frac{X - \bar{X}}{\sigma_X} + \frac{Y - \bar{Y}}{\sigma_Y} = \mathrm{konst}$$

$$Y = aX + b \;//$$

KARAKTERISTIČNE FUNKCIJE

⇒ po definiciji, KARAKTERISTIČNA FUNKCIJA varijable X jest:

$$\vartheta_X(t) = E[e^{itX}] = \sum_k p_k e^{itx}$$

↖ prinjet da je ova vrla bliska Fourierove transformaciji

⇒ SVOJSTVA KARAKTERISTIČNIH FUNKCIJA:

a) karakteristična funkcija jednoznačno određuje slučajne varijable (to znači da možemo raditi inverz)

b) ako su $X_1, \dots, X_m$ nezavisne varijable, tada je:

$$\boxed{\vartheta_{X_1+\dots+X_m} = \vartheta_{X_1}(t) \cdot \dots \cdot \vartheta_{X_m}(t)}$$

c) vrijedi: $E(X^m) = \frac{\vartheta^{(m)}(0)}{i^m}$

tj. $E(X) = \frac{\vartheta'(0)}{i}$ //

ZAD: Slučajna varijabla X poprima vrijednosti u skupu $\{-2, -1, 0, 1, 2\}$ s jednolikom razdiobom. Odredi njenu karakterističnu funkciju.

$$X \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

$$\varphi_X(t) = \sum_{k=1}^{5} p_k e^{itk} = \frac{1}{5}e^{-2it} + \frac{1}{5}e^{-it} + \frac{1}{5}e^{0} + \frac{1}{5}e^{it} + \frac{1}{5}e^{2it}$$

$$\cos(t) = \frac{1}{2}\left(e^{-it} + e^{it}\right)$$

$$\varphi_X(t) = \frac{2}{5}\cos(t) + \frac{2}{5}\cos(2t) + \frac{1}{5}$$

$$\varphi_X'(0) = 0$$

$$E(X) = \frac{0}{i} = 0$$

GEOMETRIJSKA RAZDIOBA

⇒ geometrijska razdioba je ponavljanje pokusa do prve realizacije nekog događaja

X = broj ponavljanja

$X \sim G(p)$; $p \rightarrow$ vjerojatnost realizacije promatranog događaja

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \dots & n & \dots \\ p & (1-p)p & (1-p)^2p & (1-p)^3p & (1-p)^4p & \dots & (1-p)^{n-1}p & \dots \end{pmatrix}$$

↳ ovaj zakon razdiobe jest ujedno i definicija geometrijske razdiobe

⇒ OČEKIVANJE: $E(X) = \sum_{n=1}^{\infty} n(1-p)^{n-1} \cdot p = p \cdot \frac{1}{(1-(1-p))^2} = \frac{1}{p}$

⇒ DISPERZIJA: $D(X) = \sum_{n=1}^{\infty} n^2(1-p)^{n-1}p - \left(\frac{1}{p}\right)$

$\vdots$

$$D(X) = \frac{1}{p^2} - \frac{1}{p}$$

⇒ KARAKTERISTIČNA FUNKCIJA:

$$\varphi_x(t) = \sum_{m=1}^{\infty} e^{itm}(1-p)^{m-1} \cdot p = pe^{it} \cdot \sum_{m=1}^{\infty} \left(e^{it}(1-p)\right)^{m-1}$$

$$\varphi_x(t) = \frac{pe^{it}}{1-(1-p)e^{it}}$$

↳ za vježbu, provjeri da pomoću derivacije karakteristične funkcije stvarno možemo dobiti očekivanje:

$$E(x) = \frac{\varphi'(0)}{i} = \frac{1}{p}$$

ZAD: Iz mape od 52 karte izvlačimo jednu po jednu s vraćanjem dok ne izvučemo asa ili kartu boje trefa. Neka je X broj izvlačenja. Izračunaj vjerojatnost da je broj izvlačenja veći od očekivanog broja izvlačenja.

$$P(X > E(X)) = ?$$

$$p = \frac{16}{52} = \frac{4}{13}$$

$$E(X) = \frac{13}{4} = 3.25$$

$$P(X > 3.25) = ?$$

$$P(X > 3.25) = (1-p)^3 \quad \longrightarrow \text{gledamo suprotno!}$$

$$P(X) = \left(\frac{9}{13}\right)^3$$

⇒ KORISNO SVOJSTVO:

$$\boxed{P(X > k) = (1-p)^k}$$

„k" → broj ponavljanja („ajde broj")

⇒ TEOREM ODSUDSTVA PAMĆENJA:

↳ varijabla X ima geometrijsku razdiobu ako i samo ako vrijedi:

$$\boxed{P(X = k+m \mid X > k) = P(X = m)} \qquad \forall k,m \geq 1$$

↳ „ako mi u 5 bacanja nijednom nije pala šestica, onda mi je sada vjerojatnost da dobijem 6 veća"

↳ takvo razmišljanje je pogrešno jer geometrijska vjerojatnost ne „pamti" prošle događaje

↳ DOKAZ:

- ako X ima geometrijsku razdiobu, onda vrijedi odsutstvo pamćenja

$$P(X = k+m \mid X > k) = \frac{P(X = k+m)}{P(X > k)} = \frac{(1-p)^{k+m-1} \cdot p}{(1-p)^k}$$

$$P(X = k+m \mid X > k) = P(X = m)$$

ZAD: U kutiji je 1 bijelo i 4 crna kuglica. Izvlačimo jednu po jednu s vraćanjem dok ne izvučemo bijelu.

a) Odredi: $E(X) = ?$

$P(X \leq 15 \mid X > 10) = ?$

$$X \sim G\left(\frac{1}{5}\right) \Rightarrow E(X) = \frac{1}{p} = 5$$

$$P(X \leq 15 \mid X > 10) = \text{„}\begin{matrix}\text{odsutstvo}\\ \text{pamćenja}\end{matrix}\text{“} = P(X \leq 5)$$

$P(X \leq 5) \Rightarrow$ možemo izračunati na dva načina

$$P(X \leq 5) = P(X=1) + \ldots + P(X=5)$$

$$P(X \leq 5) = 1 - P(X > 5) = 1 - \left(\frac{4}{5}\right)^5 = 0.6723$$

b) Sve je ista, al izvlačimo dok drugi put ne izvučemo bijelu! Odredi $E(X)$!

$$X \sim \begin{pmatrix} 2 & 3 & 4 & \ldots & n & \ldots \\ \frac{1}{5} & \frac{1}{5} & \binom{2}{1}\frac{4}{5}\left(\frac{1}{5}\right)^2 & \binom{3}{2}\left(\frac{4}{5}\right)^2\left(\frac{1}{5}\right)^2 & (n-1)\left(\frac{4}{5}\right)^{n-2}\left(\frac{1}{5}\right)^2 & \end{pmatrix}$$

$$E(X) = \sum_{n=2}^{\infty} n(n-1)\left(\frac{4}{5}\right)^{n-2}\left(\frac{1}{5}\right)^2 \rightarrow \text{dva puta derivacija geometrijskog reda}$$

$$E(X) = 10$$

BINOMNA RAZDIOBA

⇒ javlja se kada:

- ponavljamo pokus „n" puta, a X bilježi koliko se puta realizirao neki događaj

$$X \sim B(n, p)$$

„n" → broj ponavljanja pokusa

„p" → vjerojatnost da se traženi događaj ostvari

$$\boxed{P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}} \qquad k = 0, 1, 2, \ldots, n$$

⇒ OČEKIVANJE:

$$E(X) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

↳ ovo je komplicirana suma, pa ćemo očekivanje dobiti preko karakteristične funkcije

⇒ KARAKTERISTIČNA FUNKCIJA:

$$\varphi_X(t) = \sum_{k=0}^{n} e^{itk} \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

$$\varphi_X(t) = \sum_{k=0}^{n} \binom{n}{k} (e^{it} p)^k \cdot (1-p)^{n-k}$$

↳ binomna formula:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$$\varphi_X(t) = (p \cdot e^{it} + 1 - p)^n$$

⇒ OČEKIVANJE:

$$E(X) = \frac{\varphi'(0)}{i}$$

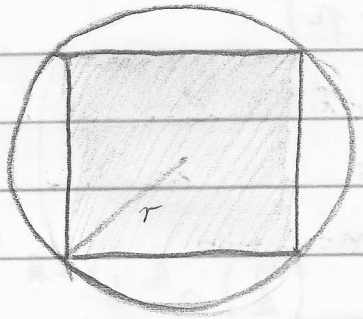
$$\varphi'(t) = n(p e^{it} + 1 - p) = p e^{it} \cdot i$$

$$E(X) = np$$

⇒ DISPERZIJA: (preko druge derivacije karakteristične funkcije)

$$D(X) = np(1-p) = npq \qquad q = 1-p$$

ZAD: U krug je upisan kvadrat. Izračunaj vjerojatnost da od 10 slučajno odabranih točaka unutar kruga, baren dvije leže unutar kvadrata?



$r = a\sqrt{2}/2$

$p = \frac{a^2}{r^2\pi}$

$$p = \frac{a^2}{\frac{a^2 \cdot 2}{4}\pi} = \frac{2}{\pi}$$

↳ ovo je binomna razdioba: $B(10, \frac{2}{\pi})$

$$P(X \geq 2) = 1 - P(X=1) - P(X=0)$$

$$= 1 - \binom{10}{1} \cdot \frac{2}{\pi} \cdot \left(1 - \frac{2}{\pi}\right)^9 - \left(1 - \frac{2}{\pi}\right)^{10}$$

$$= 0.99$$

⇒ STABILNOST BINOMNE RAZDIOBE:

↳ za neravisne X_1 i X_2 vrijedi:

$$\left.\begin{array}{l} X_1 \sim B(n_1, p) \\ X_2 \sim B(n_2, p) \end{array}\right\} X_1 + X_2 \sim B(n_1+n_2, p)$$

↳ „svejedno je da l kocku bacamo 15 puta za redom il bacamo 10 puta, odmorimo se pa još bacimo još 5 puta"

↳ **DOKAZ:**

$$\vartheta_{X_1+X_2}(t) = \vartheta_{X_1}(t) \cdot \vartheta_{X_2}(t) = (pe^{it}+1-p)^{n_1} \cdot (pe^{it}+1-p)^{n_2}$$

$$\vartheta_{X_1+X_2}(t) = (pe^{it}+1-p)^{n_1+n_2} \;//$$

⇒ BERNOULIJEVA (indikatorska) SLUČAJNA VARIJABLA:

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix} \qquad \begin{array}{l} E(X_i) = p \\ D(X_i) = p(1-p) \end{array}$$

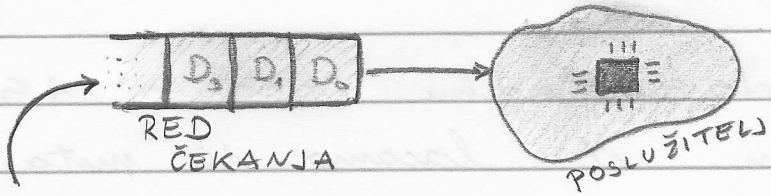
↳ tada je $X_1 + X_2 + \ldots + X_n \sim B(n, p)$

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = np \;//$$

$$D(X) = D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D(X_i) = np(1-p) \;//$$

POISSONOVA RAZDIOBA

⇒ opisuje intenzitet pojavljivanja nečega



DOLASCI

vrijeme na poslužitelju se modelira EKSPONENCIJALNOM RAZDIOBOM (o tome kasnije)

↳ upravo te dolaske se modelira Poissonovom razdiobom ⇒ λ

$$X \sim \mathcal{P}(\lambda)$$

PRIMJER 10 min, 300 jedinica je ušlo u sustav ▽.

$$\lambda = 30 \text{ zahtjev/min}$$

⇒ izračun vjerojatnosti:

$$\boxed{P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}} \qquad k = 0, 1, 2 \ldots$$

$\lambda \Rightarrow$ intenzitet dolazaka

⇒ disperzija: $D(X) = \lambda$ ⇒ bitno svojstvo

⇒ očekivanje:

$$E(x) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

$$E(x) = \sum_{k=0}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda}$$

$$E(x) = \lambda e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

↳ McLaurinov razvoj eksponencijalne funkcije:

$$e^x = \sum_{k=0}^{\infty} \frac{x^m}{m!}$$

$$E(x) = \lambda e^{-\lambda} \cdot e^{\lambda}$$

$$E(x) = \lambda$$

⇒ karakteristična funkcija:

$$\varphi_x(t) = \sum_{k=0}^{\infty} e^{itk} \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it} \cdot \lambda)^k}{k!} = e^{-\lambda} \cdot e^{e^{it} \cdot \lambda}$$

$$\varphi_x(t) = e^{\lambda(e^{it}-1)}$$

ZAD: Na neki server u prosjeku dolazi 180 mailova po satu. Izračunaj vjerojatnost da u jednoj minuti na server dođu barem tri maila?

↳ izrazimo lambdu u jedinici u kojoj se traži vjerojatnost:

$$\lambda = \frac{180}{h} = 3/min$$

$$P(X \geq 3) = 1 - P(X=2) - P(X=1) - P(X=0) =$$

$$= 1 - \frac{3^2}{2!} e^{-3} - \frac{3^1}{1!} e^{-3} - \frac{3^0}{0!} e^{-3} =$$

$$= 0.5768$$

ZADO: Na benzinsku postaju u prosjeku stiže 40 automobila po satu. Ako postoji samo jednu pumpu, auto treba jedno minuta od dolaska do odlaska, kolika je vjerojatnost da se pojavi red pred pumpom?

$$\lambda = \frac{40}{h} = \frac{2}{3} \text{ /min}$$

$$P(X \geq 2) = 1 - P(X=1) - P(X=0) =$$

$$= 1 - \frac{\lambda}{1!} e^{-\lambda} - \frac{\lambda^0}{0!} e^{-\lambda} =$$

$$= 0.144$$

b) Kolika je vjerojatnost da se pojavi red ako benzinska postoji postavi još jednu pumpu?

$$P(X \geq 3) = 1 - P(X=2) - P(X=1) - P(X=0) =$$

$$= 1 - \frac{\lambda^2}{2!} e^{-\lambda} - \frac{\lambda}{1!} e^{-\lambda} - e^{-\lambda} =$$

$$= 0.03$$

⇒ SVOJSTVA POISSONOVE RAZDIOBE 9

- **STABILNOST**

↳ možemo zbrojiti parametre (λ), ako imamo dva neravisna pohusa

$$\left.\begin{array}{l} X_1 \sim \mathcal{P}(\lambda_1) \\ X_2 \sim \mathcal{P}(\lambda_2) \end{array}\right\} \text{neravisne} \Rightarrow X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$$

↳ **dokaz:**

$$\vartheta_{X_1+X_2}(t) \overset{\text{neovisnost}}{=} \vartheta_{X_1}(t) \cdot \vartheta_{X_2}(t) =$$

$$= e^{\lambda_1(e^{it}-1)} \cdot e^{\lambda_2(e^{it}-1)} =$$

$$= e^{(\lambda_1+\lambda_2)(e^{it}-1)}$$

//

- **APROKSIMACIJA BINOMNE RAZDIOBE POISSONOM**

↳ neka je „n" velik, a „p" malen

↳ tada vrijedi:

$$B(n,p) \approx \mathcal{P}(\lambda)$$

$$\boxed{\lambda = n \cdot p}$$

↳ pokušajmo ovo sada dokazati

↳ **dokaz:**

• uvrstimo $p = \frac{\lambda}{n}$ u binomnu razdiobu

$$\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{1}{k!}\cdot\frac{n(n-1)(n-2)\cdot\ldots\cdot(n-k+1)}{n^k}\lambda^k\cdot\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!}\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right)\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-k} =$$

↳ pustimo limes: $\lim_{n\to\infty}$

↳ svi članovi idu u nulu, jedino znamo za limes:

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

$$= \frac{\lambda^k}{k!}e^{-\lambda}$$

↳ ovime je dokaz gotov

↳ ova aproksimacija je pogerno bila bitna za računanje binomnih koeficijenata sa velikim brojevima (bez kalkulatora)

ZAD: Stroj proizvodi 99 % ispravnih proizvoda. Odredi vjerojatnost da su u uzorku od 300 barem 4 proizvoda neispravna!

$$B(n=300,\ p=0.01) \approx P(n \cdot p) = P(3)$$

↳ možemo aproksimirati jer vidimo da je „n" velik, a „p" malen

$$P(X \geq 4) = 1 - P(X=3) - P(X=2) - P(X=1) - P(X=0) =$$

$$= 1 - \frac{3^3}{3!} e^{-3} - \frac{3^2}{2!} e^{-3} - \frac{3}{1} e^{-3} - e^{-3} =$$

$$= 0.35$$

⇒ mogli smo računati odmah pomoću binomne razdiobe:

$$1 - \binom{300}{3} \cdot 0.01^3 \cdot 0.99^{297} - \binom{300}{2} 0.01^2 0.99^{298} - \ldots\ldots$$

↳ primjeti da je ovo znatno kompliciranije

ZAD: Enterprise je ispalio 1000 projektila na neprijateljski brod. Vjerojatnost pogotka svakog projektila je 0.005. Ako je brod pogođen, vjerojatnost da je neutraliziran je 0.05. Aproksimirajući Poissonovom razdiobom, izračunaj vjerojatnost da će brod biti neutraliziran!

$$\left.\begin{aligned} n &= 1000 \\ p &= 0.005 \cdot 0.05 \end{aligned}\right\} \lambda = n \cdot p = 0.25$$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-0.25} =$$

$$= 0.22$$

NEPREKINUTE SLUČAJNE VARIJABLE
(DEFINICIJE I SVOJSTVA)

PR: Biramo (na sreću) broj iz intervala $[1, 4]$.

↳ ne možemo izračunati vjerojatnost da odaberemo broj manji od 3

↳ vjerojatnost da izaberemo neki određeni broj (npr. 2) je nula

⇒ **DEFINICIJA:**

↳ preslikavanje $X : \Omega \to R$ je slučajna varijabla ako $\forall x \in R$ je skup

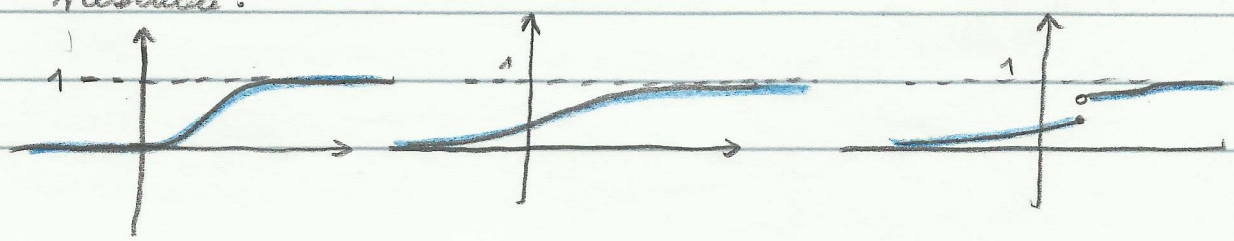
$$A_x = \{\omega \in \Omega : X(\omega) < x\}$$ događaj.

↳ funkcija razdiobe varijable X je $F : R \to [0, 1]$ definirana Δ:

$$F(x) = P(X < x)$$

BITNO! ▽

↳ pretpostavljamo da je funkcija razdiobe uvijek rastuća:



⇒ BITNA SVOJSTVA:

a) $P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$

b) F je rastuća funkcija

↳ ovo direktno slijedi jer je funkcija razdiobe njezogatnost, a njezogatnost je monotona ...

c) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$

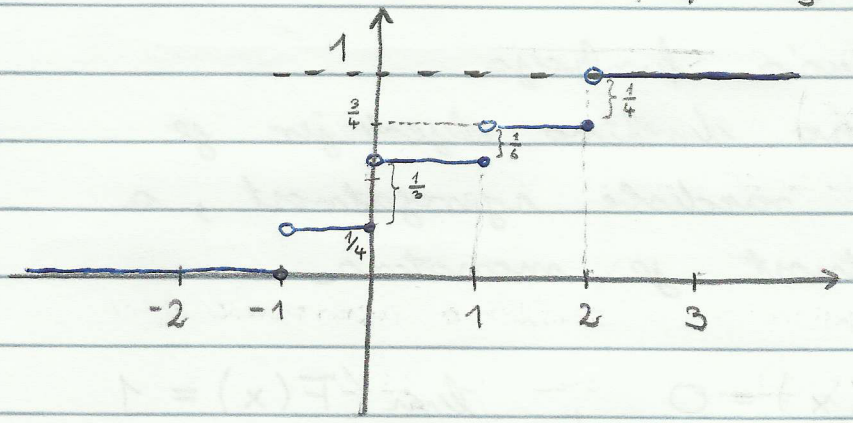
d) F je neprekinuta funkcija s lijeva

↳ ovo slijedi iz definicije:

npr. $P(X < 3) = F(3)$

ZAD: Odredite i skicirajte funkciju razdiobe slučajne varijable:

$$X \sim \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{6} & \frac{1}{4} \end{pmatrix}$$



$$F(-1) = P(X < -1) = 0$$

$$F(-0.9) = F(X < -0.9) = \frac{1}{4}$$

$$F(0) = P(X < 0) = \frac{1}{4}$$

$$F(1) = P(X < 1) = \frac{1}{4} + \frac{1}{3}$$

$$F(3) = P(X < 3) = 1$$

↳ ovo je primjer stepenaste funkcije razdiobe, a to bas i nije dobro jer ovakvu funkciju teško derivramo (samo po dijelovima)

⇒ DEFINICIJA:

• za slučajnu varijablu X kažemo da je **neprekinuta** ako postoji nenegativna funkcija $f: R \to R^+$ takva da je:

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

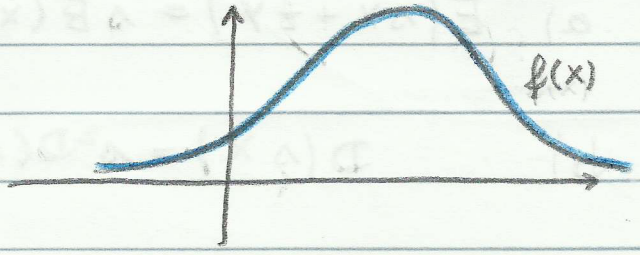
$f(x)$ ⇒ funkcija gustoće vjerojatnosti

⇒ primjetimo da vrijedi:

$$F'(x) = f(x)$$

⇒ SVOJSTVA FUNKCIJE GUSTOĆE:

↳ primjer funkcije gustoće Gaussove varijable:



$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)\,dt = \text{„površina ispod funkcije gustoće"}$$

a) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

b) $P(a \leq X < b) = \int_{a}^{b} f(x)\,dx$

BITNO!

$P(a \leq x \leq b) =$

$= P(a < X < b) =$

$= P(a \leq X < b) =$

$= P(a < X \leq b)$

↳ jer je:

$P(X = a) = 0$

⇒ OČEKIVANJE:

$$E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

→ vrijedi analogija diskretnih slučajnih varijabli

⇒ DISPERZIJA:

↳ analogno diskretnom slučajnom varijablama:

$$D(x) = \int_{-\infty}^{\infty} x^2 f(x)\, dx - E(x)^2$$

⇒ SVOJSTVA: (dokaz su analogni kao kod diskretnih $\Sigma \to \int$)

a) $E(\lambda X + t Y) = \lambda E(x) + t E(Y)$

b) $D(\lambda X) = \lambda^2 D(x)$

c) ako su X i Y nezavisne varijable, tada vrijedi:

- $E(XY) = E(x) \cdot E(y)$
- $D(X+Y) = D(x) + D(Y)$

⇒ KARAKTERISTIČNA FUNKCIJA:

$$\varphi_x(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

BITNO!

ZAD: Zadana je funkcija gustoće:

$$f(x) = \begin{cases} x^2, & x \in [0,1> \\ 2x - x^2, & x \in [1,2> \end{cases}$$

Odredi očekivanje te vjerojatnost $P(\frac{1}{2} < x < \frac{3}{2})$!

$$E(x) = \int_0^1 x \cdot x^2 dx + \int_1^2 x(2x - x^2) dx = \frac{7}{6}$$

$$P(\tfrac{1}{2} < x < \tfrac{3}{2}) = \int_{1/2}^1 x^2 dx + \int_1^{3/2} (2x - x^2) dx = \frac{3}{4}$$

↳ odredi i funkciju razdiobe!

$F(x) = ?$

Za $x \in [0,1>$:

$$F_1(x) = \int_0^x t^2 dt = \frac{t^3}{3}\Big|_0^x = \frac{x^3}{3}$$

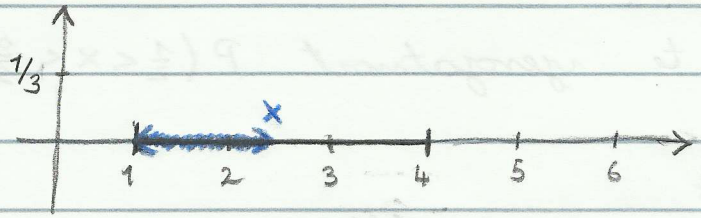
Za $x \in [1,2>$:

$$F_2(x) = \int_{-\infty}^x f(t)\, dt = \int_{-\infty}^0 0\, dt + \int_0^1 t^2 dt + \int_1^x (2t - t^2) dt$$

$$F_2(x) = -\frac{x^3}{3} + x^2 - \frac{1}{3}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{3}, & x \in [0,1> \\ -\frac{x^3}{3} + x^2 - \frac{1}{3}, & x \in [1,2> \\ 1, & x \geq 2 \end{cases}$$

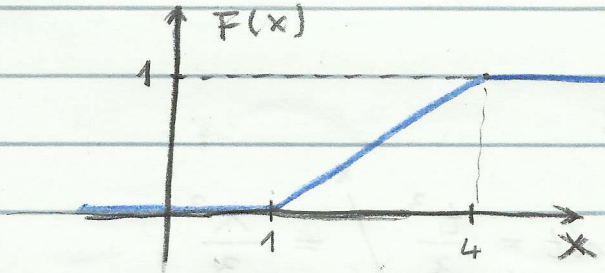
ZAD: Biramo broj iz intervala $[1, 4]$. Odredite funkciju gustoće!



↳ kada god tražimo funkciju gustoće, onda prvo računamo funkciju razdiobe:

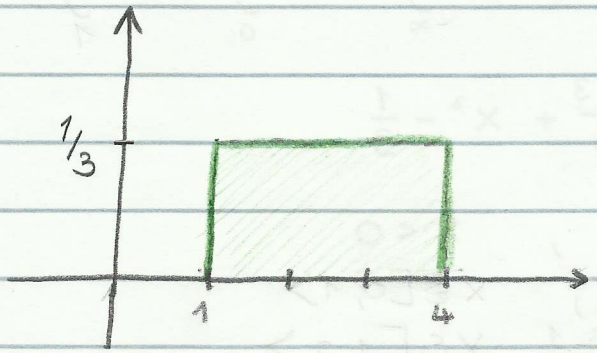
$$F(x) = P(X < x) = \text{„geometrijska vjerojatnost“}$$

$$F(x) = \frac{x-1}{3}, \quad x \in [1, 4]$$



↳ funkciju gustoće odredujemo po definiciji:

$$f(x) = F'(x) = \frac{1}{3}, \quad x \in [1, 4]$$



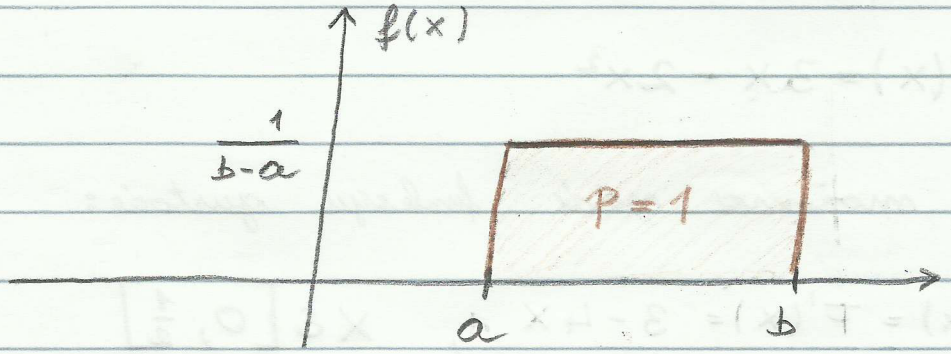
⇒ JEDNOLIKA (uniformna) RAZDIOBA:

$$X \sim U[a,b]$$

$$F(x) = \frac{x-a}{b-a} \quad \Rightarrow \quad f(x) = \frac{1}{b-a}$$

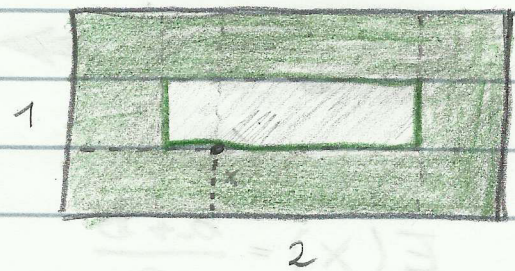
$$E(x) = \frac{a+b}{2} \qquad D(x) = \frac{(b-a)^2}{12}$$

↳ uniformna razdioba je razdioba kod koje je funkcija gustoće konstanta:



ZAD: Biramo točku unutar pravokutnika stranica 1 i 2. Neka je X udaljenost do najbliže stranice. Odredi očekivanje od X!

$$E(X) = ?$$



↳ prvo nalazimo funkciju razdiobe:

$$F(x) = P(X < x)$$

↳ zeleno su povolju ishodi za proizvoljni izbor točke

$$F(x) = \frac{2-(2-2x)(1-2x)}{2}$$

$$F(x) = 3x - 2x^2$$

↳ sada možemo naći funkciju gustoće:

$$f(x) = F'(x) = 3 - 4x \quad , \quad x \in \left[0, \tfrac{1}{2}\right]$$

bitno je dobro odrediti granice, za provjeru, možemo izračunati integral i vidjet je li on 1:

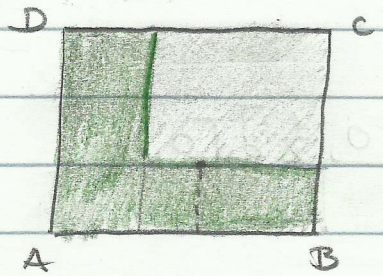
$$\int_0^{1/2} (3-4x)\,dx = 1$$

↳ sada možemo izračunati očekivanje:

$$E(x) = \int_0^{1/2} x(3-4x)\,dx$$

$$E(x) = \frac{5}{24}$$

ZAD: Biramo točku unutar kvadrata stranice jedan. Neka je X manja ili jednaka od dviju udaljenosti do stranica $\overline{AB}$, $\overline{AD}$. Odredi očekivanje od X?



$x \in [0, 1]$

$$F(x) = P(X < x) = \frac{1-(1-x)^2}{1} = 1-1+2x-x^2$$

$$F(x) = 2x - x^2, \quad x \in [0, 1]$$

↳ sada odredimo funkciju gustoće:

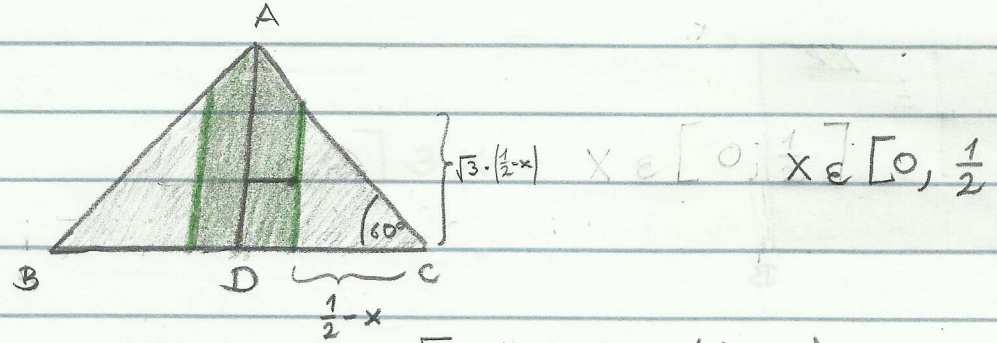
$$f(x) = F'(x) = 2-2x, \quad x \in [0, 1]$$

↳ također, lako je dobiti očekivanje iz funkcije gustoće:

$$E(x) = \int_0^1 x(2-2x)\,dx$$

$$E(x) = \frac{1}{3}$$

ZAD: Unutar jednakostraničnog trokuta ABC stranice 1 leva se točka. Neka je X udaljenost do $\overline{AD}$, ako je D polovište od $\overline{BC}$? Odredi očekivanje od X?



$X \in [0, \frac{1}{2}]$

$$F(x) = P(X < x) = \frac{\frac{\sqrt{3}}{4} - \left(\frac{1}{2}-x\right)\sqrt{3}\left(\frac{1}{2}-x\right)}{\frac{\sqrt{3}}{4}}$$

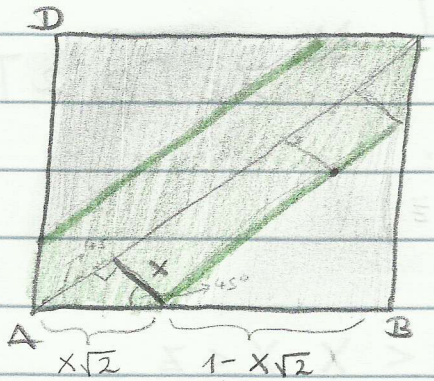
$$F(x) = 1 - 4\left(\frac{1}{2}-x\right)^2$$

$$f(x) = F'(x) = 4 - 8x \quad , \quad x \in [0, \tfrac{1}{2}]$$

$$E(X) = \int_0^{1/2} x(4-8x)\,dx$$

$$E(X) = \frac{1}{6}$$

ZAD: Biramo točku unutar kvadrata stranice 1. X je udaljenost te točke do $\overline{AC}$. Odredi očekivanje varijable X ?



$$X \in \left[0, \frac{\sqrt{2}}{2}\right]$$

$$F(x) = P(X < x) = \frac{1-(1-x\sqrt{2})^2}{1}$$

$$F(x) = \frac{1-1+2x\sqrt{2}-2x^2}{1} = 2\sqrt{2}x - 2x^2$$

↳ funkcija gustoće:

$$f(x) = F'(x) = 2\sqrt{2} - 4x \quad , \quad x \in \left[0, \frac{\sqrt{2}}{2}\right]$$

$$E(x) = \int_0^{\sqrt{2}/2} x(2\sqrt{2} - 4x)\,dx$$

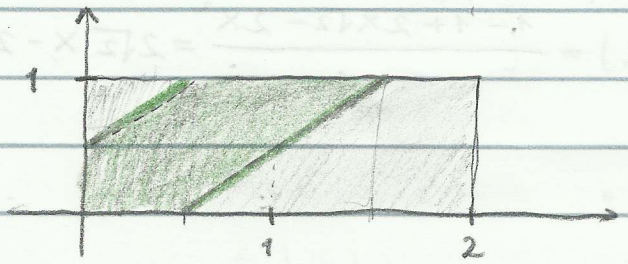
↓

ovo je jednostavno za integrirat

ZAD: Biramo dva brojá, jedan iz $[0, 2]$, drugi iz $[0, 1]$. Odredi očekivanje od Z koja je apsolutna vrijednost varijable ta dva brojá!

$$\left.\begin{array}{l} X \in [0, 2] \\ Y \in [0, 1] \end{array}\right\} \quad Z = |X - Y| \quad \Rightarrow \quad Z \in [0, 2].$$

$$F(z) = P(Z < z) = P(|X - Y| < z)$$



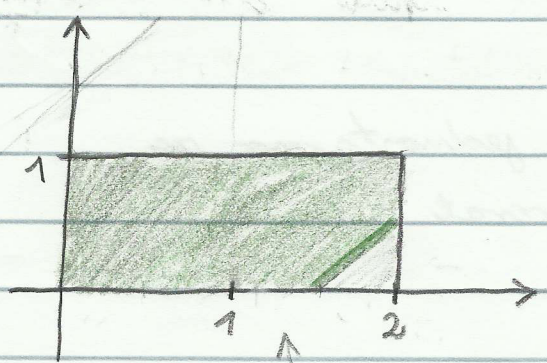
$$-z < X - Y < z$$

$$Y < X + z \qquad Y > X - z$$

skica za $z = \frac{1}{2}$

ova skica vrijedi samo za $z \in [0, 1]$

↳ nacrtajmo sada skicu za $z \in [1, 2]$:



skica za $z = 3/2$

↳ sada ćemo računati za svaki interval posebno

• interval $z \in [0, 1]$:

površina trapeza

$$F(z) = \frac{2 - \frac{1}{2}(1-z)^2 - \frac{1}{2}(2z + 2 - (1-z)) \cdot 1}{2} = z - \frac{1}{4} z^2$$

$z \in [0, 1]$

• interval $z \in [1, 2]$:

$$F(z) = \frac{2 - \frac{1}{2}(2-z)^2}{2} = 1 - \frac{1}{4}(2-z)^2$$

$z \in [1, 2]$

⇒ izračunajmo funkciju gustoće za oba intervala:

• prvi interval:

$$f(z) = 1 - \frac{1}{2} z, \quad z \in [0, 1]$$

• drugi interval:

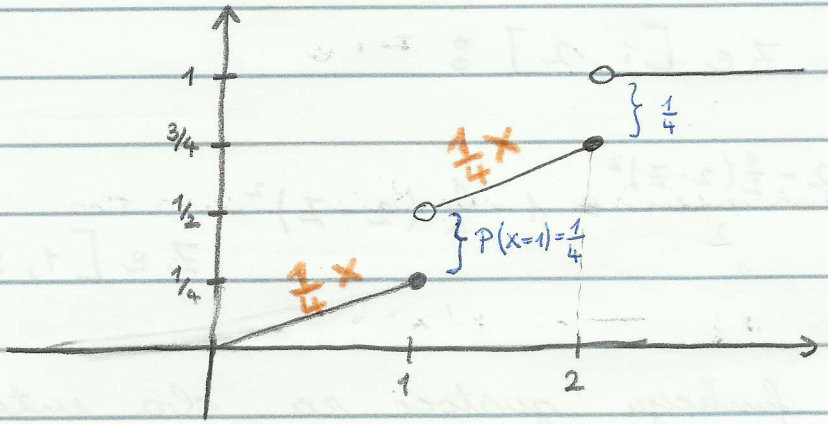
$$f(z) = 1 - \frac{1}{2} z, \quad z \in [1, 2]$$

↳ slučajna su izpale iste funkcije gustoće, pa možemo izračunati integral od 0 do 2 (da to nije izpalo isto, imali bi dva integrala):

$$E(z) = \int_0^2 z \cdot \left(1 - \frac{1}{2} z\right) dz = \frac{2}{3} //$$

RIEMANN-STIELTJESOV INTEGRAL

PR: Imamo funkcija koja izgleda ovako:



↳ ovdje ne možemo računat očekivanje na standardni način te trebamo poopćenje integrala

vrijednost
veličina skoka

$$E(X) = \int_0^1 x \cdot \frac{1}{4}\,dx + 1 \cdot \frac{1}{4} + \int_1^2 x \cdot \frac{1}{4}\,dx + 2 \cdot \frac{1}{4}$$

$$E(X) = \frac{5}{4}$$

DODATNO O KARAKTERISTIČNIM FUNKCIJAMA

- karakteristična funkcija: $\varphi_x(t) = \int_{-\infty}^{\infty} f(x)\, e^{itx}\, dx$

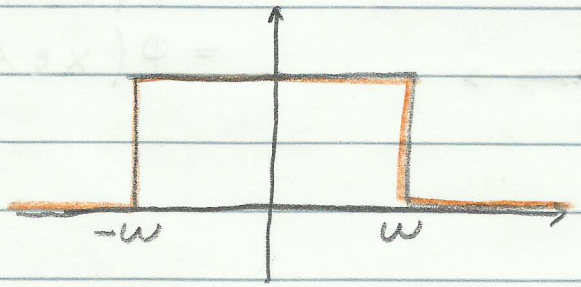
$$\varphi_x(t) = E(e^{itx})$$

PR: Odredi karakterističnu funkciju jednolike razdiobe!

$$X \sim U[a,b] \longrightarrow f(x) = \frac{1}{a-b}, \quad x \in [a,b]$$

$$\varphi_x(t) = \int_a^b \frac{1}{b-a} e^{itx}\, dx = \frac{1}{b-a} \cdot \frac{e^{itx}}{it} \Big|_a^b = \frac{e^{itb} - e^{ita}}{(b-a)it}$$

↳ konkretno za $x \in [-w, w]$:



$$\varphi_x(t) = \frac{e^{itw} - e^{-itw}}{2iwt}$$

$$\varphi_x(t) = \frac{\sin(wt)}{wt}$$

} pripremi da se ovo napravo Fourierove transformacije

FUNKCIJE SLUČAJNIH VARIJABLI

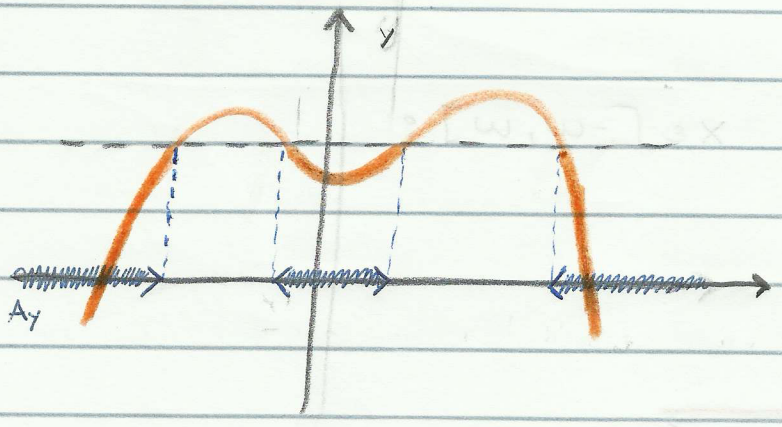
⇒ neka je $\Psi: R \to R$ realna funkcija i neka je X slučajna varijabla

↳ tada je $Y = \Psi(X)$ također slučajna varijabla

↳ zanima nas:

- $G(y)$
- $g(y)$

⇒ sagledajmo neku funkciju:



$G(y) = ?$

$G(y) = P(Y < y)$

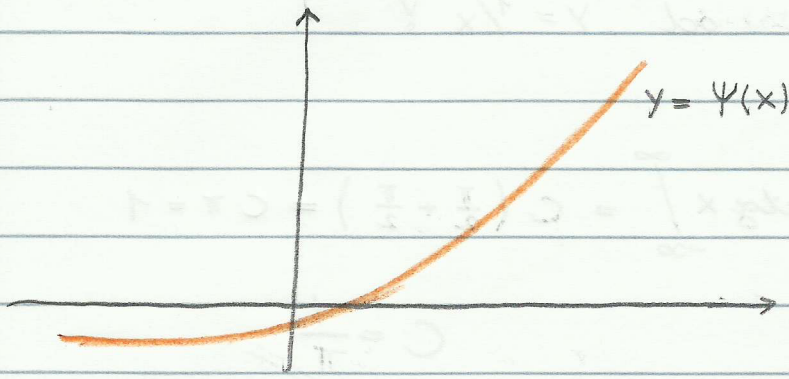
$= P(\Psi(x) < y)$

$= P(x \in A_y)$

↳ A_y je skup svih X-eva gdje je zadovoljen uvjet da je $\Psi(x) < y$

↳ mi trebamo samo izračunati vjerojatnost da se X nađe u tom skupu

⇒ neka je Ψ monotono rastuća funkcija:



↳ budući da je funkcija injektivna, možemo naći inverz dane funkcije pa možemo naći općenit postupak invertiranja računanja po intervalu Δy koji je „rastepen":

$$G(y) = P(Y < y) = P(\Psi(X) < y) =$$

$$= P(X < \Psi^{-1}(y)) = F(\Psi^{-1}(y))$$

$$\Downarrow$$

$$g(y) = G'(y) = f(\Psi^{-1}(y)) \cdot \frac{d\Psi^{-1}(y)}{dy}$$

↳ zaključak:

$$g(y) = f(x) \cdot \left| \frac{dx}{dy} \right|, \quad x = \Psi^{-1}(y)$$

→ apsolutna se javlja zbog općenitosti za monotono padajuće funkcije (njihova derivacija je negativna)

↳ ovo vrijedi samo ako je $\Psi(x)$ injekcija

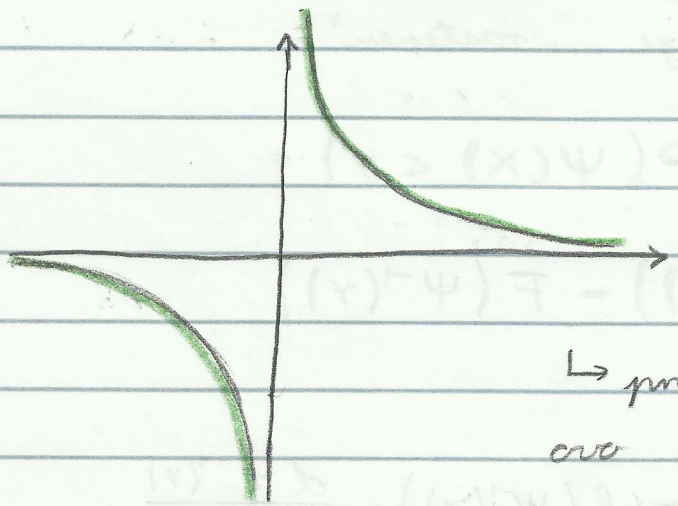
ZAD: X ima funkciju gustoće $f(x) = \frac{C}{1+x^2}$.
Odredite gustoću od $Y = 1/X$?

$$\int_{-\infty}^{\infty} \frac{C}{1+x^2}\,dx = C \operatorname{arctg} x \Big|_{-\infty}^{\infty} = C\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = C\pi = 1$$

$$C = \frac{1}{\pi}$$

↳ sada znamo funkciju gustoće:

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in R$$



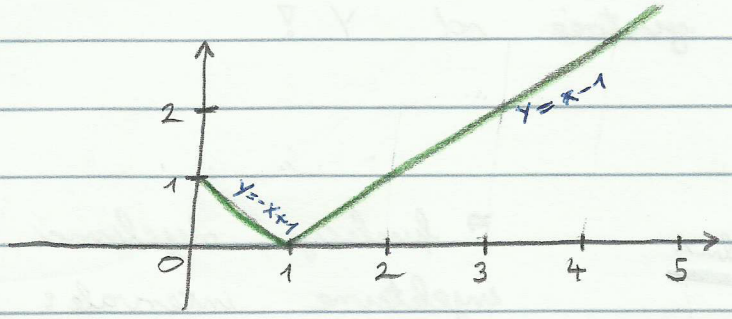
$$y = \frac{1}{x} \Rightarrow x = \frac{1}{y}$$

↳ prema slici vidimo da je ova injekcija, pa koristimo onu formulu za injektivne funkcije:

$$g(y) = \frac{1}{\pi\left(1+\left(\frac{1}{y}\right)^2\right)} \cdot \frac{1}{y^2} = \frac{1}{\pi(1+y^2)}$$

$$g(y) = \frac{1}{\pi(1+y^2)}, \quad y \in R$$

ZAD: Neka je funkcija gustoće $f(x) = e^{-x}$, $x > 0$, a Y je: $Y = |X - 1|$. Odredi $g(x)$!



↳ ovu funkciju trebamo podijeliti na injektivne intervale (ili računati po definiciji, ali je to komplicirano):

↳ injektivni interval:

- $X \in \langle 0, 1 \rangle \Rightarrow Y \in \langle 0, 1 \rangle$

$$Y = -X + 1$$
$$\Rightarrow X = 1 - Y$$
$$\left|\frac{dx}{dy}\right| = |-1| = 1$$
$$g_1(y) = e^{-(1-y)}, \quad y \in \langle 0, 1 \rangle$$

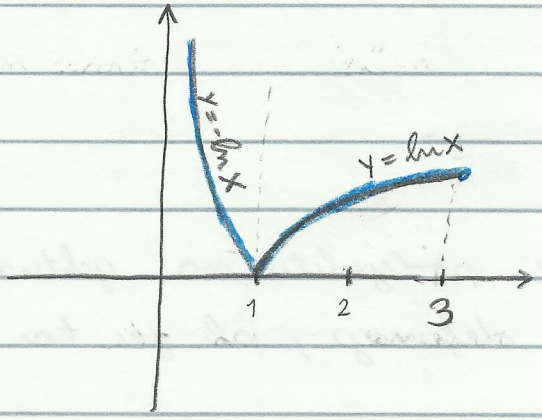
- $X \in \langle 1, +\infty \rangle$

$$Y \in \langle 0, \infty \rangle$$
$$Y = X - 1$$
$$\Rightarrow X = Y + 1$$
$$\left|\frac{dx}{dy}\right| = 1$$
$$g_2(y) = e^{-(y+1)}, \quad y \in \langle 1, \infty \rangle$$

↳ te dvije funkcije sada treba spojiti:

$$g(y) = \begin{cases} g_1 + g_2, & y \in \langle 0, 1 \rangle \\ g_2, & y \in \langle 1, \infty \rangle \end{cases}$$

ZAD: Zadana je slučajna varijabla $Y = |\ln X|$, pri čemu je X zadan s $f(X) = \frac{2}{9}X$, $X \in [0, 3]$!
Nađi funkciju gustoće od Y!



⇒ funkciju dijelimo na injektivne intervale:

a) $X \in \langle 0, 1 \rangle$

$Y \in \langle 0, \infty \rangle$

$Y = -\ln X$

$X = e^{-Y}$

$\left|\frac{dx}{dy}\right| = e^{-Y}$

$g_1(y) = \frac{2}{9} e^{-Y} e^{-Y}$

b) $X \in \langle 1, 3 \rangle$

$Y \in \langle 0, \ln 3 \rangle$

$Y = \ln X$

$X = e^{Y}$

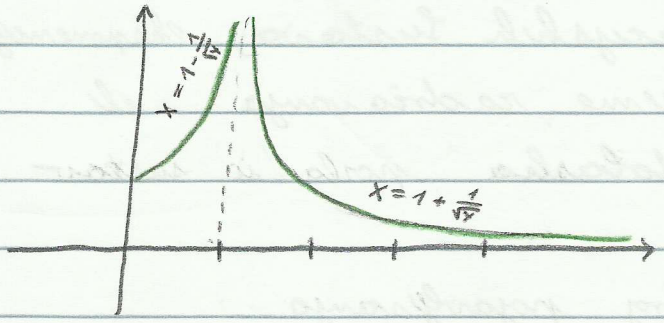
$\frac{dx}{dy} = e^{Y}$

$g_2(y) = \frac{2}{9} e^{Y} \cdot e^{Y}$

↳ sada moramo paziti na kojem intervalu se preklapaju $g_1(y)$ i $g_2(y)$:

$$g(y) = \begin{cases} g_1 + g_2, & Y \in \langle 0, \ln 3 \rangle \\ g_1, & Y \in \langle \ln 3, \infty \rangle \end{cases} = \begin{cases} \frac{2}{9} e^{-2Y} + \frac{2}{9} e^{2Y}, & Y \in \langle 0, \ln 3 \rangle \\ \frac{2}{9} e^{-2Y}, & Y \in \langle \ln 3, \infty \rangle \end{cases}$$

ZAD: X ima gustoću $f(x) = e^{-x}$, $x > 0$. Odredi gustoću od Y, ako je ona definirana ovako: $Y = \frac{1}{(X-1)^2}$ ∇



$Y = \frac{1}{(X-1)^2}$ → ovisno o kvadratnoj funkciji

$X = 1 \pm \frac{1}{\sqrt{Y}}$

⇒ ponovno dijelimo na injektivne intervale:

a) $X \in \langle 0, 1 \rangle$

$Y \in \langle 1, \infty \rangle$

$X = 1 - \frac{1}{\sqrt{Y}}$

$\frac{dx}{dy} = \frac{1}{2y\sqrt{y}}$

$g_1(y) = e^{-\left(1 - \frac{1}{\sqrt{y}}\right)} \cdot \frac{1}{2y\sqrt{y}}$

b) $X \in \langle 1, \infty \rangle$

$Y \in \langle 0, \infty \rangle$

$X = 1 + \frac{1}{\sqrt{y}}$

$\frac{dx}{dy} = \frac{-1}{2y\sqrt{y}}$

$g_2(y) = e^{-\left(1 + \frac{1}{\sqrt{y}}\right)} \cdot \frac{1}{2y\sqrt{y}}$

⇒ stoga, konačno rješenje glasi:

$$g(y) = \begin{cases} \frac{1}{2y\sqrt{y}} e^{-1 - \frac{1}{\sqrt{y}}}, & y \in \langle 0, 1 \rangle \\ \frac{1}{2y\sqrt{y}} \left(e^{-1 - \frac{1}{\sqrt{y}}} + e^{-1 + \frac{1}{\sqrt{y}}} \right), & y \in \langle 1, \infty \rangle \end{cases}$$

⇒ NAPOMENA: ako trebamo dobit funkciju razdiobe, to računamo po definiciji (POGLEDAJ TAKAV PRIMJER)

EKSPONENCIJALNA RAZDIOBA

⇒ Sjećamo se iz Operacijskih sustava, eksponencijalna razdioba modelira vrijeme raščlanjivanja ili vrijeme do prvog dolaska poziva u sustav

⇒ neka je $Z \sim \mathcal{P}(\lambda)$ broj pojavljivanja
↳ Z_x je broj pojavljivanja u intervalu $[0, x]$
↳ stoga, pošto je λ intenzitet, možemo lagano zaključiti da vrijedi:

$$Z_x \sim \mathcal{P}(\lambda x)$$

$X \Rightarrow$ vrijeme do prve pojave promatranog događaja

⇒ sada po definiciji računamo funkciju razdiobe:

$$F(x) = P(X < x) = 1 - P(Z_x = 0)$$

↳ Z je Poissonova razdioba, a za nju znamo:

$$\boxed{F(x) = 1 - e^{-\lambda x}} \qquad x > 0 \quad (\text{„}X\text{“ je vrijeme})$$

↳ funkcija gustoće:

$$f(x) = \lambda e^{-\lambda x}$$

⇒ zapis eksponencijalne razdiobe:

$$X \sim \mathcal{E}(\lambda)$$

(„λ" je isti onaj „λ" iz Poissonove razdiobe)

⇒ OČEKIVANJE:

$$E(X) = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx = \text{„parcijalna integracija"}$$

$$\boxed{E(X) = \frac{1}{\lambda}}$$

⇒ DISPERZIJA:

$$D(X) = \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} dx - \left(\frac{1}{\lambda}\right)^2 = \text{„dvostruka parcijalna integracija"}$$

$$\boxed{D(X) = \frac{1}{\lambda^2}}$$

⇒ KARAKTERISTIČNA FUNKCIJA:

$$v_X(t) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{(it-\lambda)x} dx$$

$$v_X(t) = \frac{\lambda}{it-\lambda} e^{(it-\lambda)x} \Big|_0^\infty$$

$$\boxed{v_X(t) = \frac{-\lambda}{it-\lambda}}$$

ZAD: a) Neka X ima razdiobu $X \sim \mathcal{E}(\lambda)$. Pokaži da $P(X < E(X))$ ne ovisi o λ!

$$P(X < E(X)) = P\left(X < \frac{1}{\lambda}\right) = F\left(\frac{1}{\lambda}\right)$$

$$= 1 - e^{-\lambda \cdot \frac{1}{\lambda}} = 1 - e^{-1} = 0.63 \quad //$$

b) Očekivano ispravno vrijeme rada automobila je 3 godine.

i) Odredi vjerojatnost da će se auto pokvariti tijekom prve godine!

$$P(X < 1) = 1 - e^{-\lambda X} = 1 - e^{-\frac{1}{3} \cdot 1} = 0.28 \quad //$$

ii) Pokvario tijekom treće godine!

$$P(2 < X < 3) = F(3) - F(2) = 0.14 \quad //$$

⇒ ODSUTSTVO PAMĆENJA eksponencijalne varijable:

↳ „X" ima eksponencijalnu varijablu ako i samo ako vrijedi:

$$P(X < x+t \mid X > t) = P(X < x)$$

↳ DOKAZ:

$$P(X < x+t \mid X > t) = \frac{P(t < X < x+t)}{P(X > t)} = \frac{F(x+t) - F(t)}{1 - F(t)}$$

$$= \frac{1 - e^{-\lambda(x+t)} - 1 + e^{-\lambda t}}{1 - 1 + e^{-\lambda t}} = 1 - e^{-\lambda x} =$$

$$= F(X) = P(X < x)$$

ZAD: Očekivano ispravno vrijeme rada automobila je 3 godine. Odredi vjerojatnost da se automobil neće pokvariti tijekom prve 3 godine, ako prve dvije nije bio u kvaru!

$$E(X) = 3 \quad \Rightarrow \quad \lambda = \frac{1}{3}$$

$$P(X > 3 \mid X > 2) = P(X > 1) = 1 - P(X < 1)$$

$$= 1 - \left(1 - e^{-\frac{1}{3} \cdot 1}\right) = e^{-1/3} = 0.72$$

ZAD: Vrijeme ispravnog rada računala je slučajna varijabla sa eksponencijalnom razdiobom. Vjerojatnost za ispravan rad tijekom jedne godine je 90%!

a) Odredi vjerojatnost da će ispravno raditi tijekom druge godine!

$$P(X > 1) = 0.9$$

$$1 - F(1) = 1 - (1 - e^{-\lambda \cdot 1}) = 0.9 \Rightarrow \lambda = -\ln 0.9$$

$$P(X > 2) = 1 - F(2) = 1 - (1 - e^{-(-\ln 0.9) \cdot 2}) = 0.81$$

b) Odredi vjerojatnost da će od 15 takvih računala barem 13 raditi ispravno druge godine!

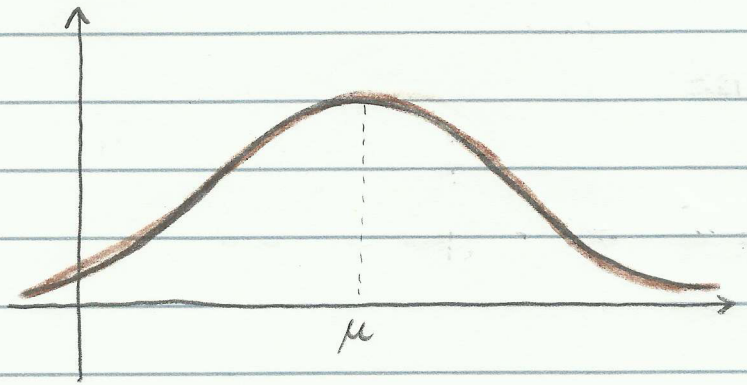
X: ono je binomna varijabla; jer ponavljamo pokus 15 puta:

$$Y \sim B(n = 15, p = 0.81)$$

$$P(Y \geq 13) = P(Y = 13) + P(Y = 14) + P(Y = 15)$$

$$= \binom{15}{13} \cdot 0.81^{13} \cdot 0.19^2 + \ldots + \binom{15}{15} \cdot 0.81^{15}$$

NORMALNA (GAUSSOVA) RAZDIOBA



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

$\mu \rightarrow$ očekivanje

$\sigma \rightarrow$ dispercija

↳ slučajna varijabla ima normalnu razdiobu ako ima gornju funkciju gustoće

- pišemo:

$$\boxed{X \sim N(\mu, \sigma^2)}$$

↳ Gaussova razdioba je zapravo ovo:

$$Y = e^{-x^2}$$

↳ integral funkcije e^{-x^2} je teško izračunat i on se aproksimira - vrijednosti tog integrala su tabelirane za različite parametre „μ" i „σ^2"

⇒ vera funkcije gustoće i tabelirane funkcije gustoće za:

$N(0, 1)$

↳ jedinična normalna varijabla

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

BITNO!

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt + \int_{0}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

$$= \frac{1}{2} + \frac{1}{2} \int_{-x}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

ova označimo Φ^* → to je tabelirana funkcija normalne varijable (vidi podsjetnik za negativne vrijednosti)

$$\boxed{\Phi(x) = \frac{1}{2} + \frac{1}{2} \Phi^*(x)}$$

⇒ formule za jediničnu normalnu varijablu:

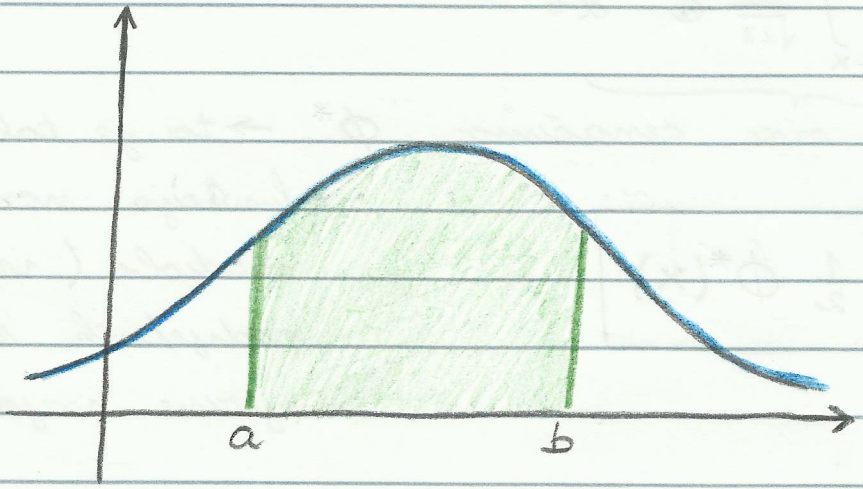
$$\boxed{P(a < X < b) = \frac{1}{2} [\Phi^*(b) - \Phi^*(a)]}$$

$$\boxed{P(X < a) = \frac{1}{2} + \frac{1}{2} \Phi^*(a)} \qquad \boxed{P(X > a) = \frac{1}{2} - \frac{1}{2} \Phi^*(a)}$$

⇒ sve ostale normalne varijable (one koje nisu jedinične) moramo centrirat i normirat i svest na jediničnu kako bi mogli koristit tablicu:

↳ to svođenje radi se pomoću formule:

$$X^* = \frac{X - \mu}{\sigma}$$



ZAD: Prosjek mečuispita je 18.37, dev = 7.22. Kolika je vjerojatnost da student ima više od 20 bodova?

$$P(X > 20) = P\left(X^* > \frac{20 - 18.37}{7.22}\right) = P(X^* > 0.226)$$

$$= \frac{1}{2} - \frac{1}{2}\,\Phi^*(0.226)$$

$$= 0.41$$

vrijednost od Φ^* vraćamo iz tablice

ZAD: Vrijeme koje student provede na putu od kuće do fakse je u prosjeku 50 min. Student ima predavanje u 8:15 i kreće u 7:20. Ako je vjerojatnost dolaska između 8:05 i 8:15 jednaka 38.3 %, kolika je vjerojatnost da će student kasnit više od 5 min?

$\mu = 50$ $\quad$ $X \rightarrow$ vrijeme na putu

$$P(45 < X < 55) = P\left(\frac{-5}{\sigma} < X^* < \frac{5}{\sigma}\right)$$

↳ vidimo da je X^* na simetričnom intervalu, pa možemo pisat:

$$P\left(-\frac{5}{\sigma} < X^* < \frac{5}{\sigma}\right) = \Phi^*\left(\frac{5}{\sigma}\right) = 0.383$$

↳ sada broj 0.383 nademo u tablici (na to je vrijednost funkcije ∇!) i dobijemo:

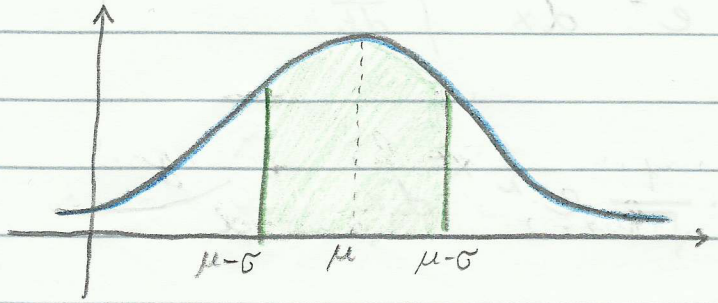
$$\frac{5}{\sigma} = \frac{1}{2} \quad \Rightarrow \quad \sigma = 10$$

↳ sada možemo izračunat što se traži:

$$P(X > 60) = P\left(X^* > \frac{60 - 50}{10}\right) = P(X^* > 1)$$

$$= \frac{1}{2} - \frac{1}{2}\,\Phi^*(1) = 0.159$$

⇒ SVOJSTVA NORMALNE RAZDIOBE:



$N(\mu, \sigma^2)$

↳ tipičan problem:

- kolko će podataka uvet ako se od očekivanja odmaknem za σ na jednu, na drugu stranu

$$P(\mu - k\sigma^2 < X < \mu + k\sigma) =$$

$$= P(-k < X^* < k) = \Phi^*(k)$$

↳ za neke brojeve „k":

$k = 1 \Rightarrow 0.68269$

$k = 2 \Rightarrow 0.9545$

$k = 3 \Rightarrow 0.9973$

$k = 4 \Rightarrow 0.99993666$

↳ iz gornjih podataka slijedi „pravilo 3 sigma":

↳ „za normalnu upotrebu dovoljno se odmaknut za 3σ od očekivanja"

⇒ KARAKTERISTIČNA FUNKCIJA:

$$\varphi_{x^*}(t) = \int_{-\infty}^{\infty} e^{itx} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \quad \Big| \frac{d}{dt}$$

$$\varphi'_{x^*}(t) = \int_{-\infty}^{\infty} e^{itx} \cdot ix \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

↳ parcijalna integracija:

$$u = e^{itx}$$

$$dv = ix \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

↳ dalje, uvodimo supstituciju:

$$z = -\frac{1}{2}x^2$$

$$\varphi'_{x^*}(t) = -t \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{itx} \cdot e^{\frac{1}{2}x^2}$$

$$\varphi'_{x^*}(t) = -t \cdot \varphi_{x^*}(t)$$ } diferencijalna jednadžba prvog reda - rješava se separacijom varijabli

$$\boxed{\varphi_{x^*}(t) = e^{-\frac{1}{2}t^2}} \Rightarrow \boxed{\varphi_x(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}}$$

↳ iz oboda računamo disperziju i očekivanje normalne varijable:

$$E(x) = \frac{\varphi'(0)}{i} \qquad D(x) = -\varphi''(0) - \mu^2$$

⇒ SVOJSTVO STABILNOSTI:

↳ neka imamo dvije neovisne varijable koje imaju normalnu razdiobu:

$$X_1 \sim N(\mu_1, \sigma_1^2)$$
$$X_2 \sim N(\mu_2, \sigma_2^2)$$

↳ tada vrijedi:

$$\boxed{s_1 X_1 + s_2 X_2 \sim N\left(s_1\mu_1 + s_2\mu_2,\ s_1^2\sigma_1^2 + s_2^2\sigma_2^2\right)}$$

↳ **DOKAZ:**

↳ preko karakteristične funkcije

$$\varphi_{s_1X_1 + s_2X_2}(t) = \varphi_{s_1X_1} \cdot \varphi_{s_2X_2} = \text{„rastavimo funkcije"}$$

↑ nezavisnost

$$= e^{it(s_1\mu_1 + s_2\mu_2) - \frac{1}{2}t^2(s_1^2\sigma_1^2 + s_2^2\sigma_2^2)}$$

(μ = $s_1\mu_1 + s_2\mu_2$, σ = $s_1^2\sigma_1^2 + s_2^2\sigma_2^2$)

ZAD: Neka su slučajne varijable zadane razdiobama:

$$X \sim N(1, 1)$$
$$Y \sim N(4, 4)$$
$$Z \sim N(9, 9)$$

Varijable su nezavisne. Odredi sljedeću vjerojatnost: $P(X \leq 3Y - 2Z)$ ∇.

$$P(X \leq 3Y - 2Z) = P(X - 3Y + 2Z \leq 0)$$

↳ zbog nezavisnosti:

$$W \sim N(1 - 3 \cdot 4 + 2 \cdot 9,\ 1 \cdot 1 + 9 \cdot 4 + 4 \cdot 9) = N(7, 73)$$

$$P(W \leq 0) = P\left(W^* \leq \frac{0-7}{\sqrt{73}}\right) = P(W^* \leq -0.819) =$$

$$= \frac{1}{2} + \frac{1}{2}\,\Phi^*(-0.819) = (*)$$

↳ Φ^* je neparna funkcija:

$$(*) = \frac{1}{2} - \frac{1}{2}\,\Phi^*(0.819) = 0.206$$

ZAD: Masa domaćih jabuka ima očekivanje 180 g i devijaciju 20 g. Masa industrijskih jabuka ima očekivanje 220 g i devijaciju 5 g. Jabučar Jan prodaje dvije domaće i dvije industrijske jabuke. Odredi vjerojatnost da je masa prodanih jabuka između 820 g i 1000 g !

$X \rightarrow$ masa „pakeća" prodanih jabuka

↳ nije dobro pisati ovako:

$$X = 2X_1 + 2X_2$$

↳ „uzeo sam jednu domaću i pomnožio masu sa 2 i to isto za industrijsku"

↳ stoga, treba pisati ovako jer linearna izvođenja:

$$X = X_1 + X_1 + X_2 + X_2$$

$$1 \cdot \sigma_1^2 \quad 1 \cdot \sigma_1^2 \quad 1 \cdot \sigma_2^2 \quad 1 \cdot \sigma_2^2$$

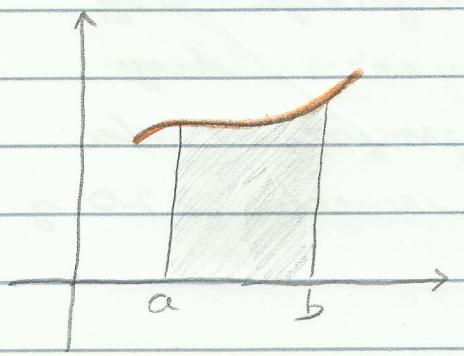
$$X \sim N(800, 850) \Rightarrow \sigma = \sqrt{850}$$

$$P(820 < X < 1000) = P(0.686 < X^* < 6.86) =$$

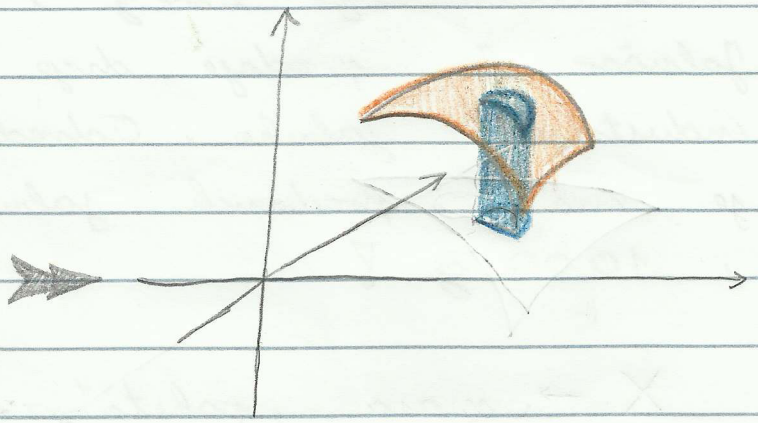
$$= \frac{1}{2}\left[\Phi^*(6.86) - \Phi^*(0.686)\right]$$

$$= 0.246$$

DVOSTRUKI INTEGRALI



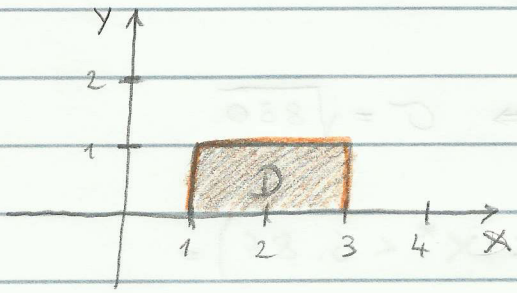
$$\int_a^b f(x)\,dx$$



$$\iint f(x,y)\,dx\,dy$$

↳ dvostruki integral je zapravo volumen

PR: Imamo funkciju, sagledajmo dvostruki integral!



$$\int_1^3 dx \int_0^1 (xy + x^2)\,dy = \int_1^3 \left(x\frac{y^2}{2} + x^2 y\right)\Big|_0^1 dx$$

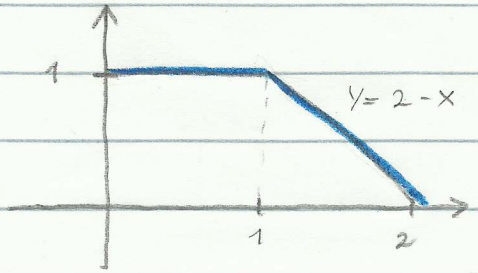
$$= \int_1^3 \left(\frac{1}{2}x + x^2\right) dx = \ldots\ldots \text{ dalje je lagano}$$

↳ no, vrijedi i suprotan poredak:

$$\int_0^1 dy \int_1^3 (xy + x^2)\,dx = \int_0^1 \left(\frac{x^2}{2}y + \frac{x^3}{3}\right)\Big|_1^3 dy = \ldots\ldots$$

dalje je lagano i dobije se isti rezultat

PR: Imamo sljedeću funkciju te moramo njen dvostruki integral!



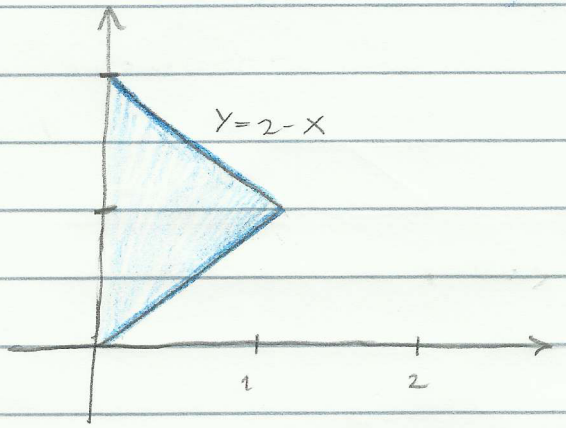
↳ postavimo sada integral, no moramo ga rastavit na dva dijela jer je područje integracije malo složenije:

$$\int_0^1 dx \int_0^1 f(x,y)\,dy + \int_1^2 dx \int_0^{2-x} f(x,y)\,dy = \text{... dugi postupak ...}$$

↳ pokušajmo suprotan poredak - možda dobijemo lakši integral:

$$\int_0^1 dy \int_0^{2-y} f(x,y)\,dx = \dots$$

PR: Imamo funkcaju i odredimo njezin dvostruki integral!



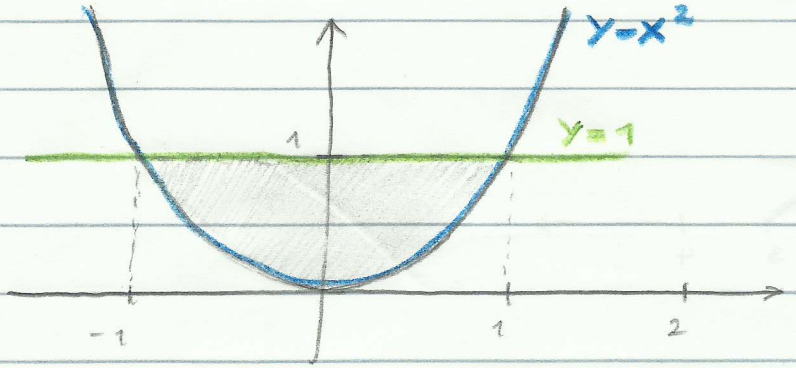
↳ prvi poredak:

$$\int_0^1 dx \int_x^{2-x} f(x,y)\,dy = \ldots \text{ lakši ishod} \ldots$$

↳ drugi poredak:

$$\int_0^1 dy \int_0^y f(x,y)\,dx + \int_1^2 dy \int_0^{2-y} f(x,y)\,dx = \ldots \text{ teži ishod} \ldots$$

PR: Odredi dvostruki integral funkcije sa slike!



↳ rešenje je:

$f(x,y) = xy$

$$\int_{-1}^{1} dx \int_{x^2}^{1} xy\,dy = \int_{-1}^{1} x \frac{y^2}{2}\Big|_{x^2}^{1} dx = \frac{1}{2}\int_{-1}^{1} x(1-x^4)\,dx =$$

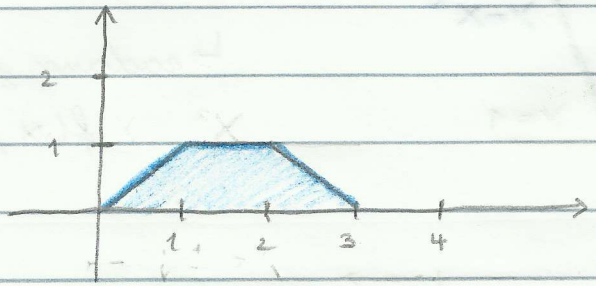
$$= \frac{1}{2}\left(\frac{x^2}{2} - \frac{x^6}{6}\right)\Big|_{-1}^{1} = 0$$

↳ drugi poredak:

$$\int_{0}^{1} dy \int_{-\sqrt{y}}^{\sqrt{y}} xy\,dx = \int_{0}^{1} y \cdot \frac{x^2}{2}\Big|_{-\sqrt{y}}^{\sqrt{y}} dy = 0$$

⇒ NAPOMENA: sarjetuje se da uvijek na prvi integral u granice stavljamo konstante

PR: Odred dvostruk integra funkcje sa slke ?



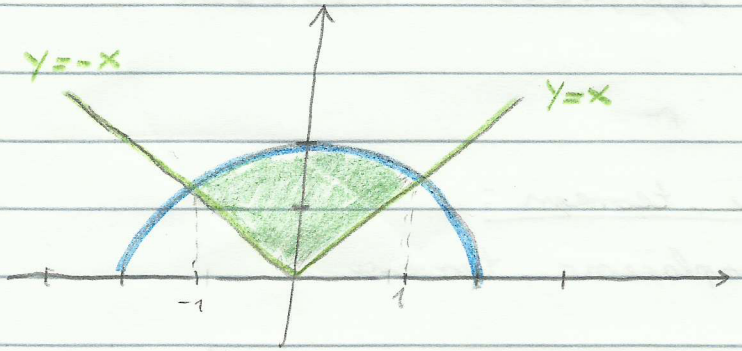
↳ odmah vidimo da je lakše integrirati u „drugom" poretku:

$$\int_0^1 dy \int_y^{3-y} f(x,y)\,dx$$

⇒ NAPOMENA: dvostruk integral bez podintegralne funkcije jest ravnina površina

$$\boxed{P = \iint dx\,dy}$$

ZAD: Odredi dvostruki integral funkcije sa slike!



$$x^2 + y^2 = 2$$

$$y = \pm\sqrt{2 - x^2}$$

↳ postavimo integral:

$$\int_{-1}^{0} dx \int_{-x}^{\sqrt{2-x^2}} f(x,y)\,dy + \int_{0}^{1} dx \int_{x}^{\sqrt{2-x^2}} f(x,y)\,dy = \text{... dalje integriramo...}$$

↳ suprotan poredak:

$$\int_{0}^{1} dy \int_{-y}^{y} f(x,y)\,dx + \int_{1}^{\sqrt{2}} dy \int_{-\sqrt{2-y^2}}^{\sqrt{2-y^2}} f(x,y)\,dx = \text{... dalje integriramo...}$$

APROKSIMACIJA BINOMNE RAZDIOBE NORMALNOM RAZDIOBOM

⇒ 1. centralni granični teorem:

↳ Moivre - Laplaceov teorem

$$X \sim B(n,p) \approx N(np, npq)$$, za dovolno velik „n"

↳ aproksimacija je to bolja što je vjerojatnost binomne varijable bliže 0.5, tj. što je razdjela „simetričnija"

ZAD: Bacamo kocku 6000 puta. Odredi vjerojatnost da smo dobili ukupno 950 i 1050 šestica?

↳ vidimo da je ovo binomna varijabla, i da je očekivan broj šestica jednak 1000

$$X \sim B\left(n = 6000, p = \frac{1}{6}\right)$$

$$P(950 < X < 1050) = \sum_{k=951}^{1049} \binom{6000}{k} \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{6000-k} = \text{„calc"} = 0.914$$

↳ naravno, trebalo je naći lakši način računanja ovakve vjerojatnosti, pa se je aproksimiralo normalnom varijablom:

$$X \sim B\left(6000, \frac{1}{6}\right) \approx N\left(1000, 833.\dot{3}\right)$$

$$P(950 < X < 1050) = P(-1.73 < X^* < 1.73) = \Phi^*(1.73) = 0.916$$

ZAD: Na ispitu od 30 pitanja imamo 2 ponuđena odgovora. Odred vjerojatnost, ako gađamo na sreću, da smo na barem 16 pitanja odgovoril točno!

$$X \sim B\left(30, \tfrac{1}{2}\right)$$

↳ točna vrijednost: $P(X \geq 16) = \sum_{k=16}^{30} \binom{30}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{30-k}$

$$P(X \geq 16) = 0.4278$$

↳ aproksimacija normalnom razdiobom:

$$X \sim N(15, 7.5)$$

$$P(X \geq 16) = P\left(X^* \geq \frac{16-15}{\sqrt{7.5}}\right) = P\left(X^* \geq 0.365\right) =$$

$$= \frac{1}{2} - \frac{1}{2}\phi^*(0.365) = 0.3575$$

primjeti da je ovo sada značajna greška jer imamo mal „n" (samo 30)

↳ zbog greške, za male „n"-ove se radi korekcija na način da se interval proširi za 0.5:

korekcija

$$P(X \geq 16) \overset{c}{=} P(X > 15.5) = P(X^* > 0.18257) =$$

$$= \frac{1}{2} - \frac{1}{2}\Phi^*(0.18257) =$$

$$= 0.4267$$

↑ puno bolja aproksimacija

⇒ <u>lokalni teorem Moivre-Laplace:</u>

$$P(X=k) = P\left(k-\frac{1}{2} < X < k+\frac{1}{2}\right)$$

↳ teorem koji govori o odnosu od vrijednosti koje želimo izračunati jer treba mnogo računanja

NEPREKINUTI SLUČAJNI VEKTORI

⇒ „n"- DIMENZIONALNI SLUČAJNI VEKTOR:

- uređena „n"- torka slučajnih varijabli

$$X = (X_1, X_2, \dots, X_n)$$

⇒ funkcija razdiobe se definira ovako:

$$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$$

⇒ mi ćemo se radržat samo na dvodimenzionalnom slučajnom vektoru:

$$\boxed{F(x, y) = P(X < x, Y < y)}$$

⇒ za vektor (X, Y) kažemo da je neprekinut, ako postoji neka nenegativna funkcija $f = R^2 \to [0, \infty>$ takva da vrijedi:

$$\boxed{F(X, Y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dx\, dy}$$

⇒ $f(x,y)$ nazivamo funkcija gustoće i vrijedi da za nju vrijedi:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \, \partial y}$$

⇒ vjerojatnost sada možemo dobiti po sljedećoj formuli:

$$P(X<x, Y<y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\, dx\, dy$$

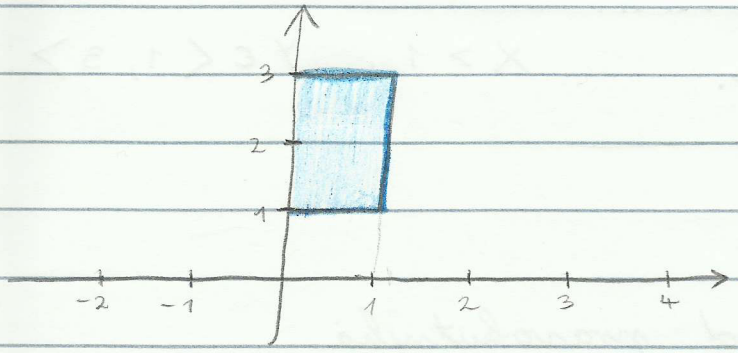
↳ definiramo i neko područje G za koje vrijedi $X, Y \in G$:

$$P(X,Y \in G) = \iint_{G} f(x,y)\, dx\, dy$$

⇒ vidimo da vrijedi:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$$

ZAD: Biramo točku unutar pravokutnika $0 \le X \le 1$, $1 \le Y \le 3$. Neka su X i Y koordinate točke. Odredi funkciju razdiobe vektora (X, Y)?

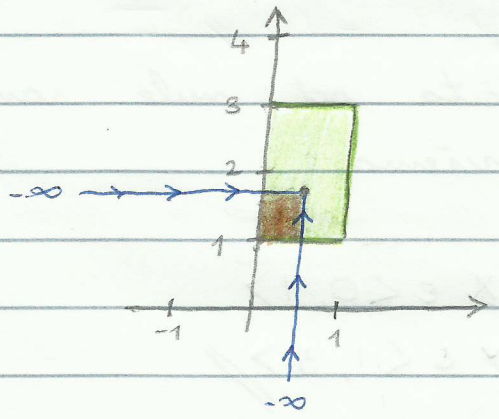


$$F(x,y) = P(X < x, Y < y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dx\,dy$$

↳ imamo 5 različitih intervala:

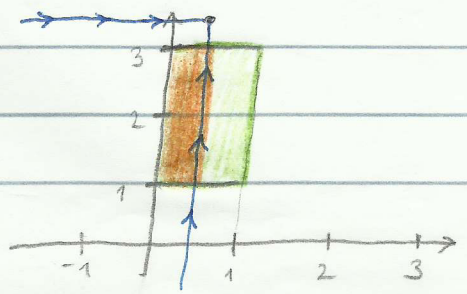
1) $F(x,y) = 0$, za $x < 0$, $\cup$ $y < 1$

2) unutar pravokutnika:



$$F(x,y) = \frac{x(y-1)}{2} \qquad x \in \langle 0,1\rangle \quad y \in \langle 1,3\rangle$$

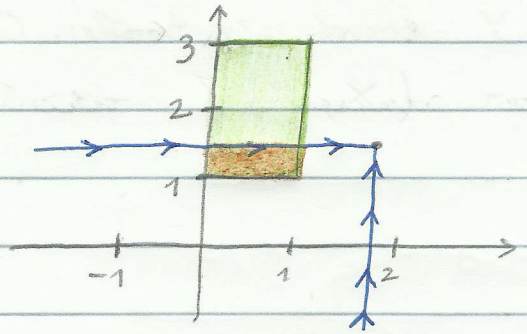
3) iznad pravokutnika:



$$F(x,y) = \frac{2 \cdot x}{2}$$

$x \in \langle 0,1\rangle$, $y > 3$

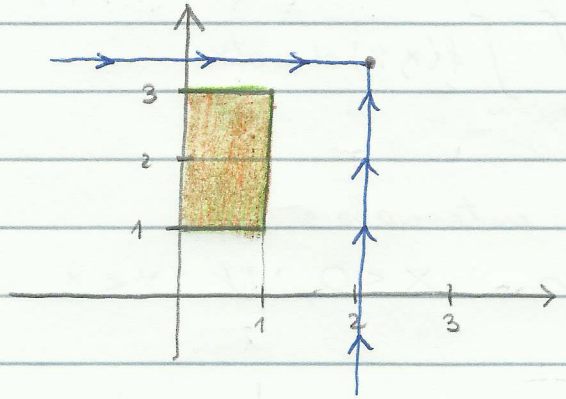
4) desna od pravokutnika



$$F(x,y) = \frac{1\cdot(y-1)}{2}$$

$x > 1, \quad y \in \langle 1, 3\rangle$

5) gore i desno od pravokutnika



$$F(x,y) = 1$$

$x > 1, \quad y > 3$

$\Rightarrow$ funkcija gustoće je različita od nule samo na području 2, pa pišemo:

$$f(x,y) = \frac{1}{2}, \qquad x \in \langle 0, 1\rangle$$
$$y \in \langle 1, 3\rangle$$

⇒ pogledajmo sljedeći problem:

- imamo zadanu razdiobu (X, Y)

↳ kako određujemo razdiobe od X i Y?

$$F_X(x) = P(X < x) = P(X < x, Y < \infty) =$$

$$= \int_{-\infty}^{x} dx \int_{-\infty}^{\infty} f(x,y)\, dy$$

↳ od prije znamo da je: $F_X(x) = \int_{-\infty}^{x} f(x)\, dx$

↳ izjednačimo ove dvije stvari i možemo zaključiti:

$$\boxed{\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\, dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\, dx \end{aligned}}$$

→ pamti ovako: „gustoća od „x"-a je integral po „y"-nu"

ovo ima smisla jer je to analogija diskretnih vektora

⇒ X i Y su slučajne varijable vektora (X, Y)

↳ X i Y su nezavisne varijable ako i samo ako vrijedi:

$$\boxed{f(x,y) = f_x(x) \cdot f_y(y)}$$

↳ DOKAZ:

$(\Rightarrow)$ $F(x,y) = P(X<x, Y<y) \overset{\text{nezavisnost}}{=} P(X<x) \cdot P(Y<y)$

$= F_x(x) \cdot F_y(y)$

↳ sada još samo deriviramo lijevu i desnu stranu

$(\Leftarrow)$ $P(X \in A, Y \in B) = \int_{X \in A} \int_{Y \in B} f(x,y)\, dx\, dy =$

$= \int_{X \in A} \int_{Y \in B} f_x(x) \cdot f_y(y)\, dx\, dy =$

$= \int_{X \in A} f_x(x)\, dx \cdot \int_{Y \in B} f_y(y)\, dy =$

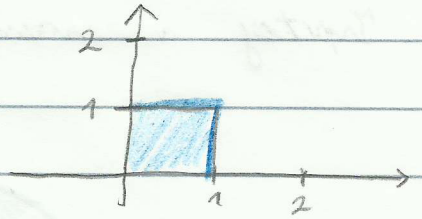
$= P(x \in A) \cdot P(y \in B)$

↓

to znači da su nezavisni po definiciji nezavisnih

ZAD: Imamo funkciju gustoće $f(X,Y) = CXY$ na $0 < X < 1$, $0 < Y < 1$.

a) Odredi konstantu C?



$$\int_0^1 dx \int_0^1 CXY\,dY = C \left.\frac{X^2}{2}\right|_0^1 \left.\frac{Y^2}{2}\right|_0^1 = \frac{C}{4} = 1$$

$$C = 4$$

b) Odredi marginalne gustoće varijabli!

$$f_X(X) = \int_0^1 4XY = 4X \left.\frac{Y^2}{2}\right|_0^1 = 2X$$

$$f_Y(X) = \int_0^1 4XY = 4Y \left.\frac{X^2}{2}\right|_0^1 = 2Y$$

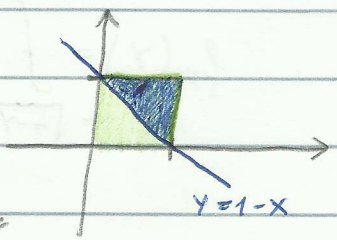
c) Jesu li X i Y nezavisni?

Jesu, jer je: $f(X,Y) = f_X(X) \cdot f_Y(Y)$

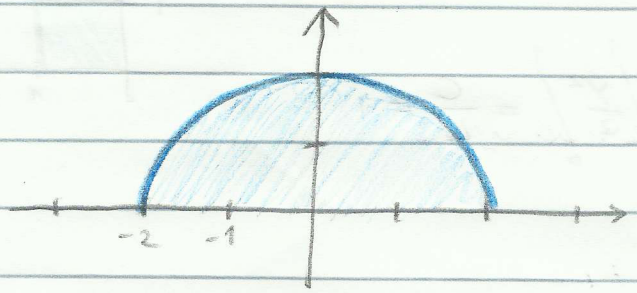
d) Odredi vjerojatnost da je zbroj $X+Y$ veći od 1!

$$P(X+Y>1) = P(Y>1-X) = \iint 4XY\,dX\,dY$$

$$= \int_0^1 dX \int_{1-X}^1 4XY\,dY = \int_0^1 4X \left.\frac{Y^2}{2}\right|_{1-X}^1 dX = \ldots = \frac{5}{6}$$



ZAD: Slučajni vektor je $f(x,y) = C$, za $x^2 + y^2 \leq 4$, $y \geq 0$. Odredi konstantu C, marginalne razdiobe. Ispitaj nezavisnost. Odredi očekivanja od X i Y.



$$\int_{-2}^{2} dx \int_{0}^{\sqrt{4-x^2}} C\, dy = C \int_{-2}^{2} \sqrt{4-x^2}\, dx = \ldots = 2\pi \cdot C = 1$$

$$C = \frac{1}{2\pi}$$

↳ odredimo marginalne razdiobe:

$$f_X(x) = \int_{0}^{\sqrt{4-x^2}} \frac{1}{2\pi}\, dy = \frac{1}{2\pi}\sqrt{4-x^2}, \quad x \in \langle -2, 2 \rangle$$

$$f_Y(y) = \int_{?}^{?} \frac{1}{2\pi}\, dx = ?$$

$$x^2 + y^2 = 4 \Rightarrow x = \pm\sqrt{4-y^2}$$

$$f_Y(y) = \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} \frac{1}{2\pi}\, dx = \frac{1}{\pi}\cdot\sqrt{4-y^2}, \quad y \in \langle 0, 2 \rangle$$

↳ očita je da su ove varijable X, Y nezavisne jer nam $f(x) \cdot f(y)$ ne može dat funkciju gustoće koja je konstanta //

↳ odredimo očekivanje:

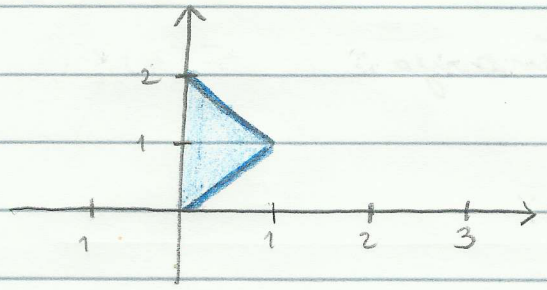
$$E(Y) = \int y f_Y(y)\, dy$$

$$E(Y) = \int_0^2 y \cdot \frac{1}{\pi} \cdot \sqrt{4-y^2}\, dy$$

$$\vdots$$

$$E = \frac{8}{3\pi} //$$

ZAD: (X, Y) ima jednoliku razdiobu na trokutu određenom sa točkama $(0,0), (1,1), (0,2)$.
Odredi marginalne gustoće i ispitaj nezavisnost.
Traženoj vjerojatnosti: $P\left(X < \frac{1}{2} \mid Y < 1\right) = ?$



↳ znamo da je funkcija gustoće kod jednolike razdiobe definirana ovako:

$$f(x,y) = \frac{1}{P_\Delta} = \frac{1}{1} = 1$$

↳ odredimo marginalne gustoće:

$$f_X(x) = \int_x^{2-x} 1 \, dy = 2 - 2x, \quad x \in \langle 0, 1 \rangle$$

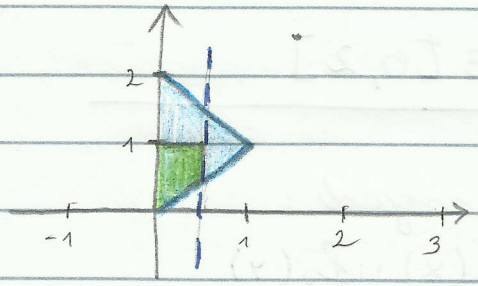
$$f_Y(y) = \int_0^{y} 1 \, dx = y, \quad y \in \langle 0, 1 \rangle$$

$$f_Y(y) = \int_1^{2-y} 1 \, dx = 2 - y, \quad y \in \langle 1, 2 \rangle$$

$$f_Y(y) = \begin{cases} y, & y \in \langle 0, 1 \rangle \\ 2 - y, & y \in \langle 1, 2 \rangle \end{cases}$$

↳ ravnost je očita jer $f(x) \cdot f(y) \neq f(x,y)$ što se vidi iz toga što je $f(y)$ radana po intervalima, pa se moramo u umnošku

↳ odredimo vjerojatnost $P(x < \frac{1}{2} \mid y < 1)$:



$$P(x < \frac{1}{2} \mid y < 1) = \frac{P(x < \frac{1}{2} \cap y < 1)}{P(y < 1)} = \frac{\int_0^{1/2} dx \int_x^1 1\, dy}{\int_0^1 dx \int_x^1 dy} = \dots$$

↳ u unutarnjem integralu imamo konstantu, pa je ovo jednako omjeru površina jer dvostruki integral konstante je upravo površina; pa odje ne trebamo računati dvostruko integrale (iako nisu teški):

$$P(x < \frac{1}{2} \mid y < 1) = \frac{\frac{3}{2} \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{3}{4} //$$

ZAD: $X \sim E$ s očekivanjem $\frac{1}{2}$, a Y ima jednoliku razdiobu na intervalu $[0, 2]$. Ako su „X" i „Y" nezavisne, odredi vjerojatnost da Y poprimi manju vrijednost od X?

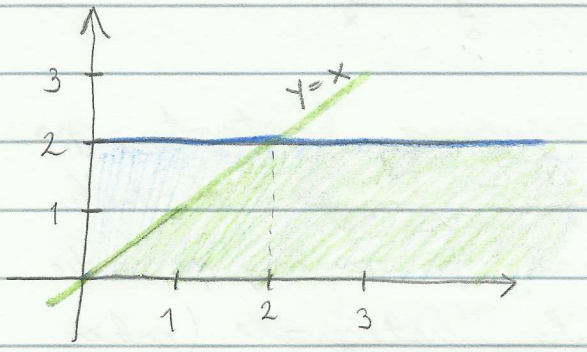
$$f_x(x) = \lambda e^{-\lambda x} = 2e^{-2x}, \quad x > 0$$

$$f_y(y) = \frac{1}{2}, \quad y \in [0, 2]$$

↳ zbog nezavisnosti vrijedi:

$$f(x, y) = f_x(x) \cdot f_y(y)$$

$$f(x, y) = e^{-2x}$$



$$P(Y < X) = \int\int e^{-2x}\,dx\,dy$$

$$P(Y < X) = \int_0^2 dy \int_y^\infty e^{-2x}\,dx$$

$$P(Y < X) = \int_0^2 \frac{e^{-2x}}{-2}\Big|_y^\infty dy$$

$$P(Y < X) = \frac{1}{2} \cdot \frac{e^{-2y}}{-2}\Big|_0^2$$

$$P(Y < X) = \frac{1}{4} - \frac{1}{4}e^{-4}$$

⇒ OČEKIVANJE VEKTORA:

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy$$

↳ drugi način zapisa: $E(\Psi(x,y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Psi(x,y)\, f(x,y)\,dx\,dy$

⇒ SVOJSTVA OČEKIVANJA:

a) $E(sX + tY) = sE(x) + tE(y)$

b) ako su X i Y nezavisne, tada je $E(x,y) = E(x)E(y)$

↳ **DOKAZ:**

a) $E(sX+tY) = \iint_{\Omega} (sX+tY)\, f(x,y)\,dx\,dy =$

$$= s\iint_{\Omega} x\, f(x,y)\,dx\,dy + t\iint_{\Omega} y\, f(x,y)\,dx\,dy =$$

$$= s\int_{-\infty}^{\infty} x\,dx \underbrace{\int_{-\infty}^{\infty} f(x,y)\,dy}_{f_x(x)} + t\int_{-\infty}^{\infty} y\,dy \underbrace{\int_{-\infty}^{\infty} f(x,y)\,dx}_{f_y(y)} = sE(x)+tE(y)$$

b) $E(XY) = \iint_{\Omega} xy\, f(x,y)\,dx\,dy = \int_{-\infty}^{\infty} x\, f_x(x)\,dx \int_{-\infty}^{\infty} y\, f_y(y)\,dy =$

$f(x,y) = f_x(x)\cdot f_y(y)$ zbog nezavisnosti

$$= E(x) + E(y)$$

UVJETNE GUSTOĆE

⇒ primjer b. lica:

↳ biramo X iz intervala $[0,1]$, a zatim Y iz intervala $[0,X]$

↳ kolika je gustoća od Y, a kolika je očekivanje od Y?

↳ ovo je vrlo čest i jednostavan problem, a još mu ne znamo postupak

⇒ neka je $f(x,y)$ gustoća vektora (X,Y), te neka je pozvata varijabla od X

↳ tada se uvjetna gustoća od Y uz uvjet $X=x$ definira ovako:

$$\boxed{f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)}}$$

} primjet da je ovo analogno formuli za uvjetnu vjerojatnost

↳ druga oblik formule koji se često koristi:

$$f(x,y) = f_{Y|X=x}(y) \cdot f_X(x)$$

↳ tada nam je marginalna gustoća od Y:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_{-\infty}^{\infty} f_{Y|X=x}(y) \cdot f_X(x)\,dx$$

$$\boxed{f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X=x}(y) \cdot f_X(x)\,dx}$$

⇒ UVJETNO OČEKIVANJE varijable „Y" koja ovisi o realizaciji varijable „X" je:

$$\boxed{E(Y|X=x) = \int_{-\infty}^{\infty} y\, f_{Y|X=x}(y)\,dy}$$

↳ tada vrijedi:

$$\boxed{E(Y) = \int_{-\infty}^{\infty} E(Y|X=x)\, f_X(x)\,dx}$$

ZAD: X je radan s $f(x) = 2x$, $x \in [0, 1]$. Biramo broj $Y \in [x, 1]$. Odredi gustoću od Y i očekivanje od Y ?

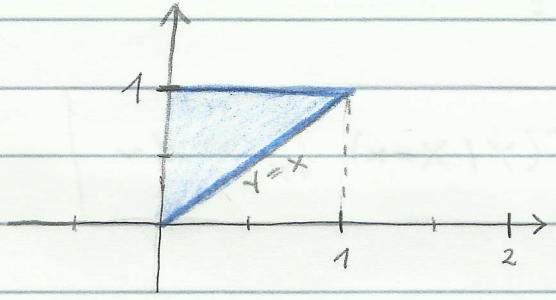
$$f_{Y|X=x}(y) = \frac{1}{1-x}, \quad y \in \langle x, 1 \rangle$$

$$f(x, y) = \frac{1}{1-x} \cdot 2x = \frac{2x}{1-x}$$

$$f_Y(y) = \int_?^? \frac{2x}{1-x}\, dx$$

↖ kako bi odredili granice, moramo nacrtati sliku:

$$x \le y \le 1$$



$$f_Y(y) = \int_0^{2y} \frac{2x}{1-x}\, dx = \ldots = -2(y + \ln|1-y|) \; ; \; y \in \langle 0, 1 \rangle$$

⇒ očekivanje možemo izračunati na dva načina:

a) prvi način - upoznali smo ga još u prvom ciklusu

- ovo može biti komplikovan integral

$$E(Y) = \int_0^1 y \cdot \left[-2(y + \ln|1-y|)\right] dy = \ldots = \frac{5}{6}$$

b) drugi način - računamo po formuli:

$$E(Y) = \int_{-\infty}^{\infty} E(Y \mid X = x)\, f_X(x)\, dx$$

- preporuča se ovaj način jer je u velikoj većini slučajeva mnogo lakši integral

$$E(Y \mid X = x) = \int_x^1 y \cdot \frac{1}{1-x}\, dy = \frac{1}{1-x} \cdot \frac{y^2}{2}\Big|_x^1 = \frac{1+x}{2}$$

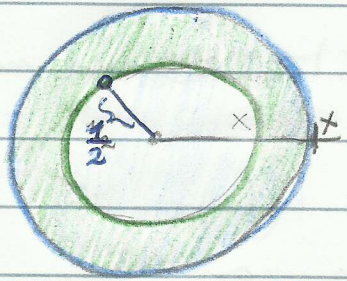
$$E(Y) = \int_0^1 \frac{1+x}{2} \cdot 2x\, dx = \frac{5}{6} //$$

⇒ vjerojatnost događaja A koji ovisi o realizaciji slučajne varijable X se definira ovako:

$$P(A) = \int_{-\infty}^{\infty} P(A \mid X = x)\, f_X(x)\, dx$$

↳ ovdje sada imamo neku klasičnu vjerojatnost koja ovisi o nekoj slučajnoj varijabli - to do sada nismo imali !!

ZAD: Radijus kruga je jednoliko distribuirana slučajna varijabla na intervalu $[1, 2]$. Biramo točku T unutar kruga. Izračunaj vjerojatnost da je točka T udaljena od središta kruga za više od $\frac{1}{2}$.



$$X \sim U[1,2] \longrightarrow f_X(x) = \frac{1}{2-1} = 1$$

$$P(A) = \int_1^2 \frac{x^2\pi - \left(\frac{1}{2}\right)^2\pi}{x^2 \cdot \pi} \cdot 1 \cdot dx$$

$$P(A) = \left(x - \frac{1}{4} \cdot \frac{x^{-1}}{-1} \right) \Bigg|_1^2$$

$$P(A) = \frac{7}{8}$$

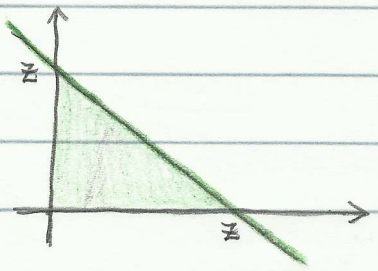
FUNKCIJE SLUČAJNIH VEKTORA

⇒ razmislimo sljedeći problem:

↳ ako znamo razdiobu od (X, Y), zanima nas razdioba od $Z = \Psi(X, Y)$

PR: (X, Y) ima funkciju gustoće $f(x, y) = e^{-x-y}$, $x > 0$, $y > 0$. Odredi funkciju gustoće i funkciju razdiobe od $Z = X + Y$?

⇒ funkciju razdiobe računamo po definiciji:



$$G(z) = P(Z < z) = P(X + Y < z)$$

$$Y < -X + z$$

$$G(z) = \iint_G f(x, y)\, dx\, dy$$

$$G(z) = \int_0^z dx \int_0^{-x+z} e^{-x-y}\, dy = \int_0^z dx\; e^{-x} \cdot (-1)\, e^{-y} \Big|_0^{-x+z} =$$

$$= \int_0^z dx\; e^{-x} \left(1 - e^{x-z}\right) = \int_0^z e^{-x} dx - \int_0^z e^{-z} dx =$$

$$= 1 - e^{-z} - z e^{-z}, \quad z \in \langle 0, \infty \rangle$$

↳ funkcija gustoće se računa kao derivacija funkcije razdiobe:

$$G'(z) = f(z) = e^{-z} - \left(e^{-z} - z e^{-z}\right) = z e^{-z}, \quad z \in \langle 0, \infty \rangle$$

⇒ prisjetimo se opće formule za funkciju gustoće slučajnih varijabli:

$$Y = \Psi(X) \quad \Rightarrow \quad g(y) = f(x) \cdot \left| \frac{dy}{dx} \right|$$

↳ postoji li neka opća formula za funkciju gustoće slučajnih vektora? Da, te ćemo tu, sada izvesti:

$$Z = \Psi(X, Y) \quad \Rightarrow \quad g(z) = ?$$

⇒ krenimo od vektora (X_1, X_2) $\quad Y_1 = \Psi(X_1)$

↳ ranimo nas (Y_1, Y_2) $\quad Y_2 = \Psi(X_2)$

↳ krećemo od vjerojatnosti:

$$P\left((X_1, X_2) \in G\right) = P\left((Y_1, Y_2) \in G\right)$$

↳ lijeva strana gornje vjerojatnosti:

$$\iint_G f(x_1, x_2)\, dx_1\, dx_2 = \left| \begin{matrix} Y_1 = \Psi(X_1) \\ Y_2 = \Psi(X_2) \end{matrix} \right| =$$

$$= \iint_G f(x_1, x_2) \left| J \right| dy_1\, dy_2$$

↳ „J" je determinanta Jakobijan:

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

↳ na desnu stranu, vjerojatnost je jednaka:

$$P\left((Y_1, Y_2) \in G\right) = \iint_G g(y_1, y_2)\, dy_1\, dy_2$$

↳ izjednačimo te dvije strane:

$$\iint_G f(x_1, x_2)\,|J|\,dy_1\,dy_2 = \iint_G g(x_1, x_2)\,dy_1\,dy_2$$

↳ odavde slijedi:

$$g(y_1, y_2) = f(x_1, x_2) \cdot \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right|$$

↳ vratimo se na početni problem kako bi pojednostavili Jakobijanovu matricu:

$$z = \Psi(x, y) \quad \longrightarrow \quad y = \chi(x, z)$$

$$\underbrace{x = x}$$

„x" se ne mijenja

↳ Jakobijan stoga iznosi:

$$\begin{vmatrix} 1 & 0 \\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial z} \end{vmatrix} = \frac{\partial y}{\partial z}$$

↳ sa pojednostavljenim Jakobijanom, formula glasi:

$$g(x, z) = f(x, y) \cdot \left| \frac{\partial y}{\partial z} \right|$$

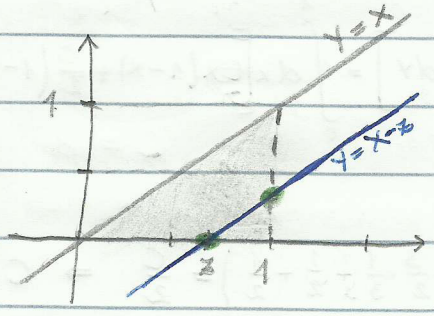
↓ iz ovoga trebamo marginalnu za z

$$\boxed{g(z) = \int_{-\infty}^{\infty} f(x, y)\,\left| \frac{\partial y}{\partial z} \right|\,dx}$$

↳ napomena: prilikom računanja uvijek uvrštavamo y !

ZAD: Zadana je funkcija gustoće $f(x,y) = 3x$, za $0 \le y \le x \le 1$. Odredi funkciju gustoće od $z = x - y$!

↳ s obzirom da se traži funkcija gustoće, možemo odmah krenut s primjenom formule:



$$y = x - z \quad , \quad \left| \frac{\partial y}{\partial z} \right| = 1$$

$$g(z) = \int_{?}^{?} 3x \cdot 1 \, dx$$

↓

granice se postavljaju u odnosu na krivulju $y = x - z$, i to za varijablu „x" (zelene točke)

a) prva točka:

$y = 0 \Rightarrow x = z$

b) druga točka:

$x = 1$

$$g(z) = \int_{z}^{1} 3x \, dx = 3 \cdot \frac{x^2}{2} \Big|_{z}^{1}$$

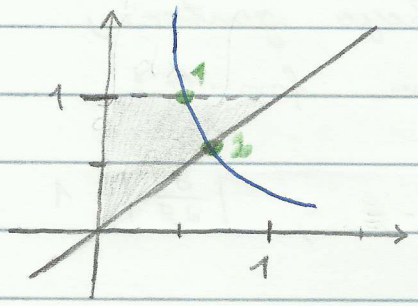
$$g(z) = \frac{3}{2}(1 - z^2) \quad , \quad z \in [0, 1]$$

moguće vrijednosti „z"-a odrečujemo iz uvjeta na varijable koje odrečuju z i same formule odrečivanja

ZAD: Oblik funkcije gustoće je dan kao:

$f(x,y) = C(x+y)$ uz uvjete $0 \le x \le y \le 1$.

Varijabla z jest dana sa $z = x \cdot y$. Odredi funkciju gustoće i očekivanje varijable z ☺



$$\int_0^1 dx \int_x^1 C(x+y)\,dy = 1$$

$$\int_0^1 dx \left[\int_x^1 Cx\,dy + C\int_x^1 y\,dy \right] = \int_0^1 dx \left[Cx(1-x) + \frac{C}{2}(1-x^2) \right] =$$

$$= C\int_0^1 \left(x - x^2 + \frac{1}{2} - \frac{1}{2}x^2 \right) dx = C\int_0^1 \left(-\frac{3}{2}x^2 + x + \frac{1}{2} \right) dx = C \cdot \left(-\frac{3}{2}\cdot\frac{1}{3} + \frac{1}{2} + \frac{1}{2} \right) = \frac{C}{2} \Rightarrow C = 2 //$$

$$f(x,y) = 2(x+y)$$

$$y = \frac{z}{x} \quad ; \quad \left| \frac{dy}{dz} \right| = \frac{1}{x} \quad ; \quad z \in [0,1]$$

↳ koristimo formulu (paziti na granice):

a) točka jedan

$$x = \frac{z}{y}$$

$$\downarrow \; y = 1$$

$$x = z$$

b) točka dva

$$y = x$$

$$x = \frac{z}{x}$$

$$x = \sqrt{z}$$

$$\left| \frac{dy}{dz} \right| = \frac{1}{x}$$

$$g(z) = \int_z^{\sqrt{z}} 2\left(x + \frac{z}{x}\right) \cdot \frac{1}{x}\,dx$$

$$g(z) = 2\int_z^{\sqrt{z}} dx + 2z\int_z^{\sqrt{z}} \frac{1}{x^2}\,dx = 2(\sqrt{z} - z) + 2z\,(-1)\,x^{-1}\Big|_z^{\sqrt{z}}$$

$$g(z) = 2(\sqrt{z} - z) + 2z\left(\frac{1}{z} - \frac{1}{\sqrt{z}}\right)$$

$$\downarrow$$

$$g(z) = 2(\sqrt{z} - z) + 2 - 2\sqrt{z} = \cancel{2\sqrt{z}} - 2z + 2 - \cancel{2\sqrt{z}}$$

$$g(z) = 2 - 2z$$

$$g(z) = 2(1 - z), \quad z \in \langle 0, 1 \rangle$$

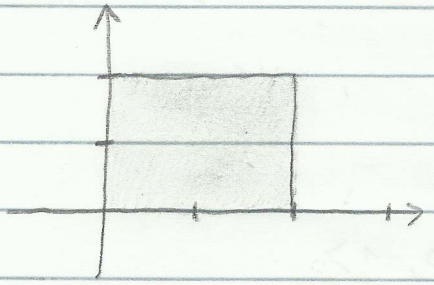
↳ očekivanje:

$$E(z) = \int_0^1 z \cdot g(z)\, dz$$

$$E(z) = 2\int_0^1 z\, dz - 2\int_0^1 z^2\, dz = 2\left(\frac{1}{2} \cdot 1^2 - \frac{1}{3} 1^3\right)$$

$$E(z) = \frac{1}{3}$$

ZAD: Zadana je funkcija gustoće $f(x,y) = C(x+y)$ $0 \le x \le 2$, $0 \le y \le 2$, a $z = y/x$. Odredi gustoću varijable z!



$$\int_0^2 dx \int_0^2 C(x+y)\,dy = 1$$

$$C\int_0^2 dx \left[x\int_0^2 dy + \int_0^2 y\,dy \right] = C\int_0^2 dx \left(2x + \frac{1}{2}\cdot 4\right) =$$

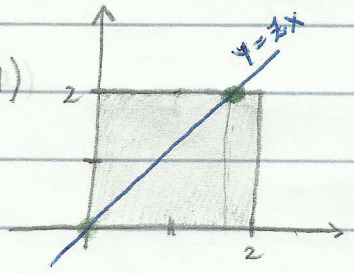
$$= 2C\int_0^2 (x+1)\,dx = 2C\left[\int_0^2 dx + \int_0^2 x\,dx\right] = 2C\left(2 + \frac{1}{2}\cdot 2^2\right) = 8C \Rightarrow C = \frac{1}{8}$$ //

↳ traži se funkcija gustoće, pa opet primjenimo sve na formulu:

$$y = x\cdot z \quad ; \quad \left|\frac{dy}{dz}\right| = x \quad ; \quad z \in \langle 0, \infty \rangle$$

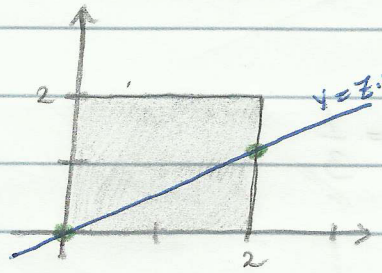
↳ ovo dijelimo na dva slučaja:

1)



$z \in \langle 1, \infty \rangle$

2)



$z \in \langle 0, 1 \rangle$

↳ na svaki od slučajeva treba sada postavit granice i izračunat funkciju gustoće

↳ ta funkcija gustoće je tada definirana po intervalima

1) a) prva točka:

$x = 0$

b) druga točka

$y = 1 \Rightarrow z = z \cdot x$

$x = \frac{1}{z}$

$$g(z) = \int_0^{2/z} \frac{x}{8}(x + zx)\,dx = \frac{1}{8}\int_0^{2/z} x^2\,dx + \frac{z}{8}\int_0^{2/z} x^2\,dx$$

$$g(z) = \frac{1}{8}(1+z)\cdot\int_0^{2/z} x^2\,dx = \frac{1}{8}(1+z)\cdot\frac{x^3}{3}\Big|_0^{2/z}$$

$$g(z) = \frac{1}{8}\cdot\frac{1}{3}(1+z)\cdot\frac{z^3}{z^3} = \frac{1+z}{3z^3} \qquad z \in \langle 1, \infty \rangle$$

2) a) prva točka:

$x = 0$

b) druga točka

$x = 2$

$$g(z) = \int_0^{2} \frac{x}{8}(x + zx)\,dx = \frac{1}{8}(1+z)\int_0^{2} x^2\,dx$$

$$g(z) = \frac{1+z}{3}, \qquad z \in \langle 0, 1 \rangle$$

↳ krajnje rješenje:

$$g(z) = \begin{cases} \frac{1+z}{3}, & z \in \langle 0, 1 \rangle \\ \frac{1+z}{3z^2}, & z \in \langle 1, \infty \rangle \end{cases}$$

ZAD: X i Y su neovisne varijable s eksponencijalnom razdiobom s očekivanjem 2. Odredi funkciju razdiobe od $Z = Y - X$ ∇

$$\lambda = \frac{1}{2}$$

↳ konstimo neovisnost:

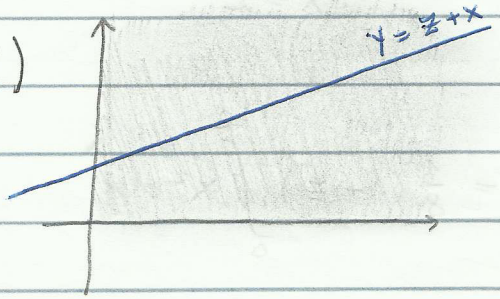
$$f(x,y) = f(x) \cdot f(y) = \frac{1}{2} e^{-\frac{1}{2}x} \cdot \frac{1}{2} e^{-\frac{1}{2}y}$$

$$f(x,y) = \frac{1}{4} e^{-\frac{1}{2}x - \frac{1}{2}y} \quad ; \quad x > 0, \; y > 0$$

↳ <u>1. NAČIN:</u> preko formule

$$y = z + x, \quad \left|\frac{\partial y}{\partial x}\right| = 1, \quad z \in \langle -\infty, \infty \rangle$$

1)



$z \in \langle 0, \infty \rangle$

$$g(z) = \int_0^\infty \frac{1}{4} e^{-\frac{1}{2}x - \frac{1}{2}(z+x)} \cdot 1 \, dx$$

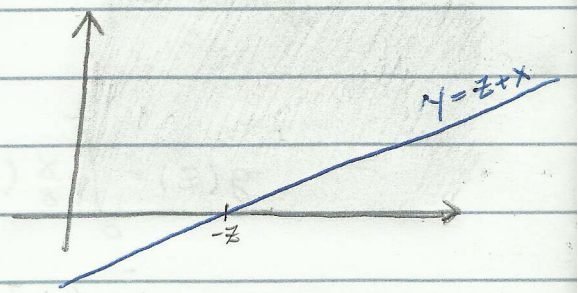
$$g(z) = \frac{1}{4} e^{-\frac{1}{2}z}$$

$$G(z) = \int_{-\infty}^{0} \frac{1}{4} e^{\frac{1}{2}z} \, dz + \int_0^z \frac{1}{4} e^{-\frac{1}{2}z} \, dz$$

$$G(z) = 1 - \frac{1}{2} e^{-\frac{1}{2}z}$$

$z \in \langle 0, \infty \rangle$

2)



$z \in \langle -\infty, 0 \rangle$

$$g(z) = \int_{-z}^\infty \frac{1}{4} e^{-\frac{1}{2}x - \frac{1}{2}(z+x)} \, dx$$

$$g(z) = \frac{1}{4} e^{\frac{1}{2}z}$$

$$G(z) = \int_{-\infty}^{z} \frac{1}{4} e^{\frac{1}{2}z} \, dz$$

$$G(z) = \frac{1}{2} e^{\frac{1}{2}z} \qquad z \in \langle -\infty, 0 \rangle$$

↳ **2. NAČIN:** preko definicije funkcije razdiobe (bolji način ako se eksplicitno traži funkcija razdiobe)

1) $$G(z) = P(Z < z) = P(Y - X < z) = \int_0^\infty dx \int_0^{x+z} \frac{1}{4} e^{-\frac{1}{2}x - \frac{1}{2}y} dy = \quad (y < x+z)$$

$$= \int_0^\infty dx \frac{1}{4} e^{-\frac{1}{2}x} (-2)\left(e^{-\frac{1}{2}(x+z)} - 1\right) = -\int_0^\infty dx \cdot \frac{1}{2} e^{-\frac{1}{2}x}\left(1 - e^{-\frac{1}{2}(x+z)}\right)$$

$$= 1 - \frac{1}{2} e^{-\frac{1}{2}z}, \quad z \in \langle 0, \infty \rangle$$

2) $$G(z) = P(Y - X < z) = \int_0^\infty dx \int_0^{x+z} \frac{1}{4} e^{-\frac{1}{2}x - \frac{1}{2}y} dy \quad (y < x+z)$$

$$G(z) = \frac{1}{2} e^{\frac{1}{2}z}, \quad z \in \langle -\infty, 0 \rangle$$

TEORIJA VJEROJATNOSTI
1 ZAKONI VELIKIH BROJEVA

⇒ ovo poglavlje je prilično teško i mnoštvo dokaza nećemo raditi jer su prekomplicirani
↳ stoga, puno ćemo shvaćati samo na intuitivnoj razini

⇒ motivacija: ponavljamo pokus „n" puta te bilježimo koliko puta se dogodio neki povoljni ishod
↳ ako se povoljan ishod dogodio „m-puta, omjer „$\frac{m}{n}$" teži vjerojatnosti događaja „p"

⇒ niz X_n <u>konvergira po vjerojatnosti</u> k slučajnoj varijabli X ako za $\forall \varepsilon > 0$ vrijedi:

$$\lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$$

↳ oznaka koja je bitna i koju ćemo sada stalno koristiti:

$$X_n \xrightarrow{P} X$$

čitaj: „konvergira po vjerojatnosti"

⇒ TEOREM: Nejednakost Markova: ako X poprima nenegativne vrijednosti, onda $\forall \varepsilon > 0$ vrijedi:

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

⇒ TEOREM: Čebiševljeva nejednakost: za svaku slučajnu varijablu X s konačnim očekivanjem i disperzijom vrijedi:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}$$

⇒ dokaz nejednakosti Markove:

$$P(X \geq \varepsilon) = \int_{\varepsilon}^{\infty} f(x)\,dx \leq \int_{\varepsilon}^{\infty} \frac{x}{\varepsilon} f(x)\,dx \leq \frac{1}{\varepsilon} \int_{0}^{\infty} x f(x)\,dx = \frac{E(X)}{\varepsilon}$$

⇒ dokaz Čebiševljeve nejednakosti:

↳ samo treba umjesto X staviti $(X - E(X))^2$, tj. treba samo centrirati varijablu

Zad. 8 Očekivanje u godini za sunčane dana je 75. Dokažite da je vjerojatnost da u godini nebude više od 200 sunčanih dana veća od $\frac{5}{8}$.

$$P(X \geq 200) \leq \frac{75}{200} = \frac{3}{8}$$

$$P(X \leq 200) \geq 1 - \frac{3}{8} = \frac{5}{8}$$

ZAD: Broj bodova na ispitu je slučajna varijabla s očekivanjem 65 i disperzijom 81. Pokaži da je vjerojatnost da je broj bodova između 55 i 75 veća od 19%.

$$E(X) = 65$$
$$D(X) = 81$$
$$P(55 < X < 75) > 0.19$$

↳ imamo interval, pa trebamo koristiti Čebiševljevu nejednakost:

$$P(|X - 65| \geq 10) \leq \frac{81}{10^2}$$

$$P(|X - 65| \geq 10) > 1 - 0.81$$

$$P(|X - 65| \geq 10) > 0.19$$

⇒ SLABI ZAKON VELIKIH BROJEVA:

$$\bar{X} = \frac{\sum X_i}{n} \xrightarrow{P} \mu$$

↳ on samo govori čemu će težit srednja vrijednost ako puno puta ponavljamo pokus

↳ ovaj zakon je samo statističke prirode

⇒ niz X_n <u>konvergira gotovo sigurno</u> k slučajnoj varijabli X ako vrijedi:

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

↳ oznaka koja je kraća i koju ćemo mi koristiti:

$$X_n \xrightarrow{g.s.} X$$

čitaj: „konvergira gotovo sigurno"

⇒ JAKI ZAKON VELIKIH BROJEVA:

$$\bar{X} = \frac{\sum X_i}{n} \xrightarrow{g.s.} \mu$$

⇒ niz X_n <u>konvergira po distribuciji</u> k varijabli X ako za odgovarajuće funkcije vrijedi:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

↳ oznaka koja je kraća, koju ćemo mi koristiti:

$$X_n \xrightarrow{D} X$$

čitaj: „konvergira po distribuciji"

⇒ TEOREM: <u>Levíjev teorem</u>:

$X_n \xrightarrow{D} X$ ako i samo ako $\varphi_{X_n}(t) \longrightarrow \varphi_X(t)$

↳ bitan teorem jer sada možemo provjeravati karakterističke funkcije umjesto samih varijabli

CENTRALNI GRANIČNI TEOREM

⇒ govori o varijabli po kojoj se raspodjeljuju rezultati slučajnog pokusa
↳ sve teži normalnoj varijabli !

⇒ TEOREM: neka je X_n niz identički distribuiranih nezavisnih slučajnih varijabli s očekivanjem „μ" i disperzijom „σ^2"
↳ tada vrijedi:

$$\frac{\sum_{k=1}^{n} X_k - n \cdot \mu}{\sigma \sqrt{n}} \xrightarrow{D} N(0,1)$$

⇒ korolar gornjeg teorema jest Moivre - Laplace teorem:

$$B(n,p) \xrightarrow{D} N(np, npq)$$

ZAD: Na ruleti je označeno 37 polja: 18 crnih 18 crnih i jedno zeleno. Igrač se kladi na crveno ili crno polje. Ako pogodi, dobije 10 kuna, ako promaši, izgubi 10 kn. Odredi vjerojatnost da nakon 100 klađenja igrač bude na dobitku!

↳ prva traženo očekivanje u jednom pojedinom klađenju:

$$X_i \sim \begin{pmatrix} -10 & 10 \\ \frac{19}{37} & \frac{18}{37} \end{pmatrix}$$

$$E(X_i) = -10 \cdot \frac{19}{37} + 10 \cdot \frac{18}{37} = -\frac{10}{37}$$

$$D(X_i) = 100 - \frac{100}{37^2}$$

↳ mi trebamo: $P\left(\sum_{i=1}^{100} X_i > 0\right) = ?$

↳ primjenimo centralni granični teorem:

$$P\left(\sum_{i=1}^{100} X_i > 0\right) = P\left(\frac{\sum X_i - 100 \cdot \left(-\frac{10}{37}\right)}{\sqrt{100 - \frac{100}{37^2}} \cdot \sqrt{100}} > \frac{0 - 100 \cdot \left(-\frac{10}{37}\right)}{\sqrt{100 - \frac{100}{37^2}} \cdot \sqrt{100}}\right)$$

$$= P(X^* > 0.27) = \frac{1}{2} - \frac{1}{2}\Phi^*(0.27)$$

$$= 0.3936$$

ZAD: Sto brojeva zaokruženo na najbliži cijeli broj i potom zbrojeno. Ako su greške jednolika distribuirane na intervalu $[-0.5, 0.5]$, nađi interval unutar kojeg će sa vjerojatnosti 95% nalaz pogreška zbroja ∇_0

$$X_i \sim U[-0.5, 0.5] \Rightarrow f(x) = \frac{1}{1} = 1$$

$$E(X_i) = 0$$

↳ trebamo i disperziju: $D(X_i) = \int_{0.5}^{0.5} x^2 f(x)\,dx - E(x)^2$

$$D(X_i) = 1/12$$

↳ sada imamo potrebne podatke, pa možemo iz na centralni granični teorem:

$$P(-a < \sum X_i < a) = 0.95$$

$$P\left(\frac{-a - 100 \cdot 0}{\sqrt{\frac{1}{12}} \cdot \sqrt{100}} < X^* < \frac{a - 0}{\sqrt{\frac{1}{12}} \cdot \sqrt{100}}\right) = \Phi^*\left(\frac{a}{\frac{10}{\sqrt{12}}}\right) = 0.95$$

$$a = 1.96 \cdot \frac{10}{\sqrt{12}}$$

$$a = 5.658$$

ZADO: Koliko puta moramo bacit kocku da bi vjerojatnost da je aritmetička sredina između 3 i 4 bila najmanje 0.95?

↳ zapišimo zadatak (zadano i traženo):

$$P(3 < \bar{X} < 4) \longrightarrow \geq 0.95$$

$$n = ?$$

↳ individualna varijanca:

$$X_i \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

$$E(X_i) = 3.5$$

$$D(X_i) = \frac{35}{12}$$

↳ sada idemo na centralni granični teorem:

$$P\left(3 < \frac{\sum X_i}{n} < 4\right) = ?$$

↓

treba prilagodit centralni granični teorem:

$$\frac{\sum X_k - n\mu}{\sigma\sqrt{n}} : n = \frac{\frac{\sum X_k}{n} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \longrightarrow N$$

$$P\left(3 < \frac{\sum X_i}{n} < 4\right) = P\left(\frac{3 - 3.5}{\sqrt{\frac{35}{12}}/\sqrt{n}} < X^* < \frac{4 - 3.5}{\sqrt{\frac{35}{12}}/\sqrt{n}}\right) =$$

$$= \Phi^*(0.2928\sqrt{n}) \geq 0.95$$

$$n = 45$$

MATEMATIČKA STATISTIKA

⇒ predmet statističkog proučavanja:
- populacija X

⇒ OBILJEŽJE ⇒ podatak koji mjerimo i proučavamo

⇒ PODJELA STATISTIKE:

a) deskriptivna statistika - analiza točnih podataka cijele populacije

b) statistička obrada uzorka - temeljem uzorka se zaključuje o cijeloj populaciji (time se mi bavimo)

⇒ UZORAK: za $(X_1, \dots, X_m)$ kažemo da je UZORAK ako su X_i međusobno nezavisni i svi su uzeti iz iste populacije X

⇒ VELIČINA (dimenzija, volumen) UZORKA:
- broj članova uzorka
- oznaka: „m"

⇒ realizacija uzorka:

$$(x_1, x_2, \dots, x_m)$$

⇒ pretpostavit ćemo da je uzorak realizacija slučajnog vektora

TOČKASTE PROCJENE

⇒ cilj: parametar populacije ϑ procijeniti na temelju realizacije uzorka $x_1, \dots, x_n$, tj. nađi funkciju $\vartheta = g(x_1, \dots, x_n)$

⇒ STATISTIKA: slučajna varijabla $\Theta = g(X_1, \dots, X_n)$

⇒ za statistiku Θ kažemo da je procjenitelj parametra ϑ, a vrijednost $\hat{\vartheta} = g(x_1, \dots, x_n)$ nazivamo PROCJENA od ϑ

⇒ KRITERIJI ZA ODABIR DOBRIH STATISTIKA:

a) za statistiku Θ kažemo da je <u>nepristrana</u> statistika parametra ϑ ako vrijedi:

$$\boxed{E(\Theta) = \vartheta}$$

b) neka su Θ_1 i Θ_2 dvije statistike
↳ kažemo da je Θ_1 bolja (efikasnija) od Θ_2 ako vrijedi $D(\Theta_1) < D(\Theta_2)$
↳ statistika je <u>valjana</u> ako $\forall \varepsilon > 0$, Θ_n konvergira prema ϑ po vjerojatnosti

⇒ BITAN TEOREM:

• nepristrana statistika je valjana ako vrijedi:

$$\boxed{D(\Theta_n) \to 0 \quad \text{kada } n \to \infty}$$

⇒ pogledajmo neke tipove statistike:

a) statistika za procjenu očekivanja:

$$\overline{X} = \frac{X_1 + \ldots + X_n}{n}$$

↳ dokaz nepristranosti:

$$E(\overline{X}) = E\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{E(X_1) + \ldots + E(X_n)}{n}$$

„a" → oznaka za očekivanje

$$E(\overline{X}) = \frac{n \cdot a}{n} = a \;//$$

↳ dokaz valjanosti:

$$D(\overline{X}) = D\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{\sum D(X_i)}{n^2}$$

$$D(\overline{X}) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0$$

b) statistika procjene disperzije uz poznato očekivanje cijele populacije:

$$D^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - a)^2$$

↳ dokaz nepristranosti:

$$E(D^2) = \frac{1}{n} \sum_{i=1}^{n} \underbrace{E(x_i - a)^2}_{D(x_i)} = \frac{1}{n} \cdot n \cdot \sigma^2 = \sigma^2$$

↳ dokaz valjanosti:

$$D(D^2) = \frac{1}{n^2} \sum_{i=1}^{n} D(x_i - a)^2 = \frac{1}{n^2} \cdot n D(x - a)^2$$

primjenimo nezavisnost

$$D(D^2) = \frac{1}{n} \left(E(x-a)^n - \left(E(x-a)^2\right)^2 \right) = \frac{\mu - \sigma^4}{n} \xrightarrow[n \to \infty]{} 0$$

c) statistika procjene disperzije ur neporanato očekivanje populacije:

$$\Theta = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

↳ dokaz nepristranost:

$$E(\Theta) = \frac{1}{n}\sum_{i=1}^{n} E(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} D(X_i - \bar{X}) =$$

$$= \frac{1}{n}\cdot\sum_{i=1}^{n} D\left(X_i - \frac{\sum X_i}{n}\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n} D\left(\frac{n-1}{n}\cdot X_i - \frac{1}{n}\sum_{j\neq i} X_j\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{n-1}{n}\right)^2 \underbrace{D(X_i)}_{\sigma^2} + \frac{1}{n^2}\cdot D\left(\sum_{i\neq j} X_j\right)\right] =$$

$$\sum_{i\neq j} \underbrace{D(X_j)}_{\sigma^2} = \Sigma$$

$$= \frac{1}{n}\left[n\cdot\left(\frac{n-1}{n}\right)^2\sigma^2 + \frac{1}{n^2}\cdot\frac{n-1}{n}\sigma^2\right] =$$

$$= \frac{n-1}{n}\sigma^2$$

→ ovo je varijanca od σ^2, pa je procjena pristrana, pa

↳ stoga, procjenu treba pomnožiti s $\frac{n}{n-1}$

$$\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

⇒ Češće koristimo još jednostavniji oblik upravo izvedene formule, pa ga i sada izvedimo:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i^2 - 2 x_i \cdot \bar{x} + \bar{x}^2 \right) =$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - 2\bar{x} \underbrace{\sum_{i=1}^{n} x_i}_{A} + \sum_{i=1}^{n} \bar{x}^2 \right)$$

$$\bar{x} = \frac{A}{n}$$

↳ odavde slijedi:

$$\boxed{S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n \cdot \bar{x}^2 \right)}$$

↑ ovo je lakše utipkati u kalkulator, pa se više koristi

ZAD: Iz intervala $[\alpha, 1]$ odabrano je „n" brojeva $X_1, \dots, X_n$. Da bi procjenili α, odaberemo statistiku $Y = \min\{X_1, \dots, X_n\} - \frac{1}{n+1}$!

a) Dokaži da statistika nije nepristrana?

b) Sa kojim faktorom treba pomnožit Y da ona postane nepristrana?

↳ dlj: $E(Y) = \alpha$; $X = \min\{X_1, \dots, X_n\}$

↳ po definiciji nađemo razdiobu, ratim to deriviramo i imamo gustoću te preko toga dobijemo očekivanje

$$F(x) = P(X < x) = P(\min(X_1, \dots, X_n) < x) =$$

$$= 1 - P(\min(X_1, \dots, X_n) > X) =$$

$$= 1 - P(X_1 > x, \dots, X_n > x) \overset{\text{nezavisnost}}{=} 1 - (P(X_i > x))^n =$$

$$= 1 - \left(\int_x^1 \frac{1}{1-\alpha} \, dx \right)^n = (*)$$

↳ za X_i imamo: $X_i \sim U[\alpha, 1]$

$$f_{X_i}(x) = \frac{1}{1-\alpha}$$

$$(*) = 1 - \left(\frac{1-x}{1-\alpha} \right)^n$$

↳ sada znamo varijablu, pa deriviranjem dobijemo gustoću:

$$f(x) = F'(x) = n\left(\frac{1-x}{1-\alpha}\right)^{n-1} \cdot \frac{1}{1-\alpha}$$

↳ iz toga računamo očekivanje:

$$E(X) = \int_{\alpha}^{1} x \cdot \frac{n(1-x)^{n-1}}{(1-\alpha)^n}\, dx = \left| \begin{matrix} 1-x=t \\ -dx=dt \\ x=1-t \end{matrix} \right| = \frac{n}{(1-\alpha)^n} \int_{0}^{1-\alpha} (1-t)\, t^{n-1}\, dt =$$

$$= \frac{n}{(1-\alpha)^n} \left[\frac{t^n}{n} - \frac{t^{n+1}}{n+1} \right] \Bigg|_0^{1-\alpha} = \frac{n}{(1-\alpha)^n} \left[\frac{(1-\alpha)^n}{n} - \frac{(1-\alpha)^{n+1}}{n+1} \right] =$$

$$= 1 - \frac{n}{n+1}(1-\alpha) \quad //$$

$$E(Y) = 1 - \frac{n}{n+1} + \frac{n\alpha}{n+1} - \frac{1}{n+1} = \frac{n\alpha}{n+1}$$

varijabla od α, pa je statistika pristrana

↳ za nepristranost, statistiku množimo sa $\frac{n+1}{n}$ kako bi $E(Y)$ bila jednaka α

ZAD: Odredi procjenu očekivanja i disperzije uzorka iz podataka?

x_i	2560	2600	2620	2650	2700
m_i	2	3	10	4	1

$$\overline{x} = \frac{2 \cdot 2560 + 3 \cdot 2600 + 10 \cdot 2620 + 4 \cdot 2650 + 1 \cdot 2700}{20}$$

$$\hat{S} = \frac{1}{20-1}\left[\sum_{i=1}^{20} m_i x^2 - n\bar{x}^2\right] = 967.4$$

KRITERIJ NAJVEĆE IZGLEDNOSTI

⇒ neka je $X_1, \dots, X_n$ realizacija uzorka populacije X čija funkcija gustoće „$f(\vartheta, X)$" ovisi o parametru „ϑ"

↳ funkcija izglednosti („likelihood function") definira se kao umnožak:

$$L(\vartheta, x_1, \dots, x_n) = f(\vartheta_1, x_1) \cdot \dots \cdot f(\vartheta_1, x_n)$$

⇒ za procjenu parametra „ϑ" uzima se izglednost ⇒ za koju funkcija „L" poprima globalni maksimum

PR: Uzorak $x_1, \dots, x_n$ uzet će se iz populacije s gustoćom $f(x) = \lambda x^{\lambda-1}$. Koristeći kriterij najveće izglednosti, odredite procjenu od „λ"!

$$L(\lambda, x_1, \dots, x_n) = f(x_1) \cdot \dots \cdot f(x_n) =$$

$$= \lambda x_1^{\lambda-1} \cdot \lambda x_2^{\lambda-1} \cdot \dots \cdot \lambda x_n^{\lambda-1} =$$

$$= \lambda^n (x_1 \cdot \dots \cdot x_n)^{\lambda-1}$$

↳ kako bi ovo lakše derivirali, možemo ovu funkciju staviti pod logaritam, pa derivirat zbroj umjesto umnoška:

- to možemo načiniti jer je logaritam injektivna i rastuća funkcija, pa neće promijeniti maksimume

$$\ln L = n \ln \lambda + (\lambda - 1) \ln (x_1 \cdot \dots \cdot x_n) \quad / \frac{d}{d\lambda}$$

$$\frac{d}{d\lambda} \ln L = \frac{n}{\lambda} + \ln (x_1 \cdot \dots \cdot x_n) = 0$$

$$\lambda = - \frac{n}{\ln (x_1 \cdot \dots \cdot x_n)}$$

$$\lambda = - \frac{n}{\sum_{i=1}^{n} \ln x_i} \longrightarrow \text{ovo je sigurno maksimum}$$

ZAD: Uzorak $X_1, \dots, X_n$ izvučen je iz populacije s gustoćom $f(x) = \frac{\lambda}{2\sqrt{x}} e^{-\lambda\sqrt{x}}$. Koristeć kriterij najveće izglednosti, odredi procjenu od „λ"!

$$L(\lambda_1, X_1, \dots, X_n) = \frac{\lambda^n}{2^n \sqrt{X_1 \cdot \ldots \cdot X_n}} e^{-\lambda(\sqrt{X_1} + \dots + \sqrt{X_n})}$$

↳ logaritmiramo:

$$\ln L = n \ln \lambda + \ln e^{-\sum_{i=1}^{n} \sqrt{X_i}} - \ln\left(2^n \sqrt{X_1 \cdot \ldots \cdot X_n}\right) \quad \Big/ \frac{d}{d\lambda}$$

$$\frac{d}{d\lambda} L = \frac{n}{\lambda} - \sum_{i=1}^{n} \sqrt{X_i} - 0 = 0$$

$$\lambda = \frac{n}{\sum_{i=1}^{n} \sqrt{X_i}}$$

ZAD: Vjerojatnost pojavljivanja pisma na nesparenom novčiću je nepoznata. Novčić je bačen 5 puta i pismo se pojavilo 2 puta. Zatim je novčić bačen 8 puta i pismo smo dobili 3 puta. Koristeći kriterij najveće izglednosti, procjeni vjerojatnost pojavljivanja pisma!

$$\left.\begin{array}{l} 5 \rightarrow 2P \\ 8 \rightarrow 3P \end{array}\right\} \quad p = \frac{5}{13} \rightarrow \text{odmah znamo rezultat}$$

↳ ovo je binomna varijabla: $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$$L(p) = P(X=2) \cdot P(X=3) = \binom{5}{2} p^2 (1-p)^3 \cdot \binom{8}{3} p^3 (1-p)^5$$

$$= \binom{5}{2}\binom{8}{3} p^5 (1-p)^8$$

↳ logaritmiramo:

$$\ln L = \ln\left[\binom{5}{2}\binom{8}{3}\right] + 5 \ln p + 8 \ln(1-p) \quad / \frac{d}{dp}$$

$$\frac{5}{p} + \frac{8}{1-p}(-1) = 0$$

$$p = \frac{5}{13}$$

ZAD: Poissonova varijabla je u tri pokusa poprimila vrijednosti 5, 7, 3. Pomoću kriterija najveće vjerodostojnosti, odredi procjenu za „λ" ?

↳ već znamo da je $\lambda = E(X) = \overline{X}$ kod Poissonove varijable, pa unaprijed znamo rezultat: $\lambda = 5$

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$L(\lambda) = P(5) \cdot P(7) \cdot P(3) = \frac{\lambda^5}{5!} e^{-\lambda} \cdot \frac{\lambda^7}{7!} e^{-\lambda} \cdot \frac{\lambda^3}{3!} e^{-\lambda} =$$

$$= \frac{\lambda^{15}}{5!\,7!\,3!} e^{-3\lambda}$$

↳ logaritmiramo:

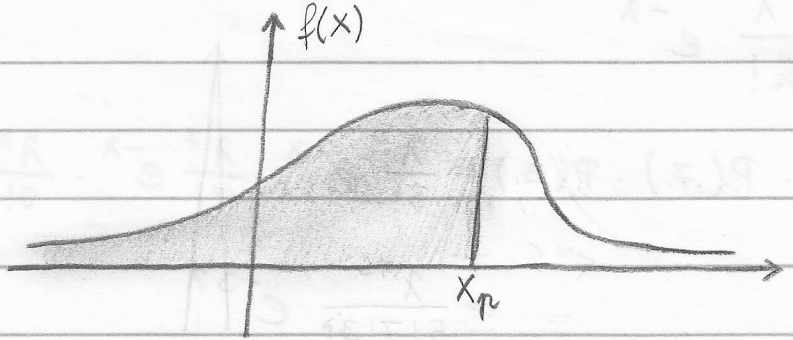
$$\ln L = 15 \ln \lambda - 3\lambda - \ln(5!\,7!\,3!) \quad / \frac{d}{d\lambda}$$

$$\frac{15}{\lambda} - 3 - 0 = 0$$

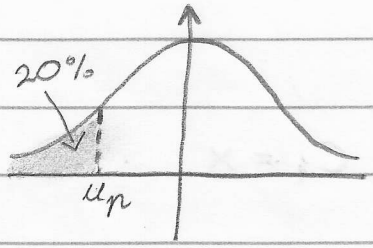
$$\lambda = 5$$

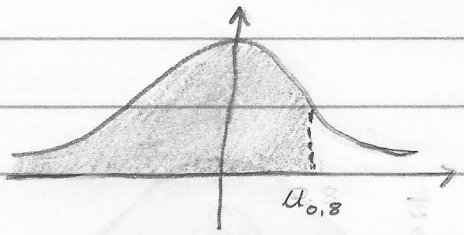
INTERVALNE PROCJENE.
KVANTILI

⇒ realan broj x_p za koji vrijedi $F(x_p) = p$,
tj. $\int_{-\infty}^{x_p} f(t)\,dt = p$, naziva se KVANTIL



ZAD: a) Koristeći tablicu 4, izračunaj kvantil „u_p" za $p = 0.2$!



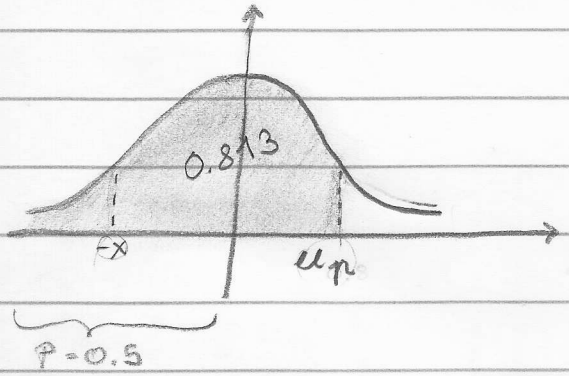


↳ tablica 4 nam govori da se radi o jediničnoj normalnoj razdiobi

$$u_{0.8} = 0.8416$$

$$u_{0.2} = -0.8416$$

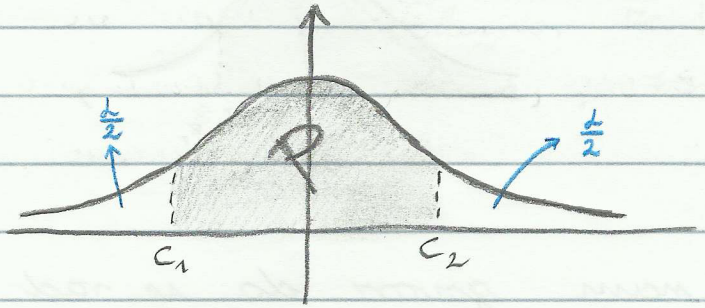
b) Odredi kvantil „u_p" za $p = 0.813$ pomoću tablice 1 !



$$(0.813 - 0.5) \cdot 2 = 0.626$$

$$u_{0.813} = 0.889$$

⇒ interval $[c_1, c_2]$ za koji vrijedi $P(c_1 < X < c_2) = p$ naziva se INTERVAL POVJERENJA reda „p" (pouzdanosti / mjera pouzdanosti „p")



$C_1 = X$

⇒ napomena - varijabla ne mora biti simetrična (kod nas će biti)

⇒ veličina $\alpha = 1 - p$ se zove NIVO ZNAČAJNOSTI (signifikantnosti)

↳ stoga, na gornjoj slici, c_1 i c_2 će nam biti kvantili:

$$C_1 = X_{\frac{\alpha}{2}} \qquad C_2 = X_{1-\frac{\alpha}{2}}$$

INTERVALNE PROCJENE ZA PARAMETRE NORMALNE RAZDIOBE

	poznat drugi parametar	nepoznat drugi parametar
za očekivanje	str. 27	str. 34
za disperziju	str. 30	str. 35

} ova su stranice algoritama za dobivanje nepoznatog

ZAD: Mjerenjem smo dobili sljedeći uzorak:

X_j	20	22	24	26	28	30	31
n_j	3	3	8	7	3	2	1

Odredi 95%-ni interval za očekivanje i 95%-ni dvostrani interval za disperziju!

$n = 27$, $\bar{X} = 25$

$$\hat{S}^2 = \frac{1}{26} \sum (n_j X_i^2 - n \cdot \bar{X}^2) = 8.846$$

$p = 0.95$, $\alpha = 0.05$

↳ uzimamo kvantile studentove razdiobe:

$$t_{26,\, 1-\frac{\alpha}{2}} = 2.056$$

$$P\left(\bar{X} - t \cdot \frac{\hat{S}}{\sqrt{n}} < a < \bar{X} + t \cdot \frac{\hat{S}}{\sqrt{n}}\right) = p$$

$$P(23.823 < a < 26.177) = 95\%$$ OKRENI →

↳ za disperziju, trebamo dvostrani interval uz neporanato očekivanje

- mi smo samo procjenili očekivanje, a ne znamo mu stvarnu vrijednost

$$\chi^2_{n-1,\,1-\frac{\alpha}{2}} = \chi^2_{26,\,0.975} = 41.923$$

$$\chi^2_{n-1,\,\frac{\alpha}{2}} = \chi^2_{26,\,0.025} = 13.844$$

$$P\left(\frac{26 \cdot 8.846}{41.923} < \sigma^2 < \frac{26 \cdot 8.846}{13.844}\right) = 0.95$$

ZAD: Imamo podatke: 16.2, 16.1, 16, 16.4, 16.5, 16.3, 16.2, 16.4, 16.2, 16.3. $\sigma = 0.3$

Za koj nivo pouzdanost je duljina intervala za očekivanje jednaka 0.3123?

$$2 \cdot u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 0.3123$$

↳ odavde dobivamo „α":

$$u_{1-\frac{\alpha}{2}} = \frac{\sqrt{3}}{\sigma} \cdot \frac{0.3123}{2}$$

$$u_{1-\frac{\alpha}{2}} = 1.645$$

↓ TABLICA 4

$$1 - \frac{\alpha}{2} = 0.95$$

$$\alpha = 0.1$$

↳ pouzdanost je $1 - \alpha$:

$$p = 1 - \alpha$$

→ ovaj „p" označuje pouzdanost (nemoj pomiješat oznake)

$$p = 90\%$$

$$P\left(\bar{x} - \frac{0.3123}{2} < a < \bar{x} + \frac{0.3123}{2}\right) = p$$

$$P(16.104 < a < 16.416) = 90\%$$

INTERVALNA PROCJENA ZA VJEROJATNOST DOGAĐAJA

ZAD: Na nekoj izlaznoj anketi od 200 glasača njih 112 glasalo je za kandidata X.

a) Odredi 95% -ni interval povjerenja za ovog kandidata!

$\alpha = 0.05$

$$p = \hat{p} \pm u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$u_{0.975} = 1.96$ (TABLICA 4)

$\hat{p} = \frac{112}{200} = 0.56$

$$p = 0.56 \pm 0.069$$

$$P(0.491 < p < 0.629) = 95\%$$

b) S kojom vjerojatnošću kandidat X može tvrditi da će sigurno biti izabran?

$$\hat{p} - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} > 0.5$$

$$0.56 - u_{1-\frac{\alpha}{2}} \sqrt{\frac{0.56 \cdot 0.44}{200}} > 0.5 \Rightarrow u_{1-\frac{\alpha}{2}} < 1.709$$

$1 - \frac{\alpha}{2} = 0.95$

$\alpha = 0.1 \Rightarrow p = 90\%$ //

c) Koliko velik vzorak ljudi moramo izpitati da bi izbor kandidata X bio siguran na nivo značajnost 95%?

$$p = 0.95 \rightarrow \alpha = 0.05$$

$$\hat{p} - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} > 0.5$$

$$0.56 - 1.96 \cdot \sqrt{\frac{0.56 \cdot 0.44}{n}} > 0.5$$

$$n > 262.93$$

$$n = 263 //$$

ZAD: Broj pristiglih automobila u jednom satu je srednje 66 automobila. Odred 90% interval zadane Poissonove razdobe!

$\bar{X} = 66$ auta/h

$\alpha = 1 - p = 0.1$

$\downarrow$

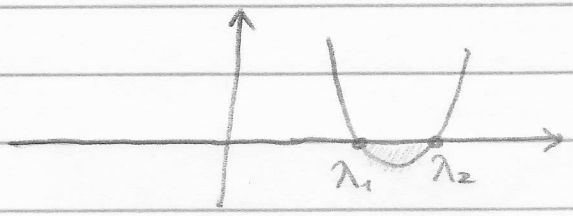
$u_{1-\frac{\alpha}{2}} = u_{0.95} = 1.645$

$$P\left(|\bar{X} - \lambda| < u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\lambda}{n}}\right) = p$$

$$|66 - \lambda| < 1.645 \cdot \sqrt{\frac{\lambda}{24}}$$

$$\lambda^2 - 132.113\lambda + 4356 < 0$$

$\lambda_1 = 63.33$, $\lambda_2 = 68.79$



$\lambda \in \langle 63.33, 68.79 \rangle$

$$P(63.33 < \lambda < 68.79) = 90\%$$

TESTIRANJE HIPOTEZA.

HIPOTEZE I POGREŠKE ODLUČIVANJA

Prs Specifikacija veli da printer ima 50 stranica u minuti. Isprobali smo 30 blokova ispisa i dobili smo 47 stranica u minuti.

$H_0 \rightarrow$ proizvođač govori istinu
(nulta hipoteza)

$H_1 \rightarrow$ proizvođač laže
(alternativna hipoteza)

$H_1: a < 50$

⇒ POGREŠKA PRVE VRSTE (α):

↳ verjetnost da smo odbrali H_1, a istina je H_0

⇒ POGREŠKA DRUGE VRSTE (β):

↳ verjetnost da smo odbrali H_0, a istina je H_1

TESTIRANJE PARAMETARSKIH HIPOTEZA

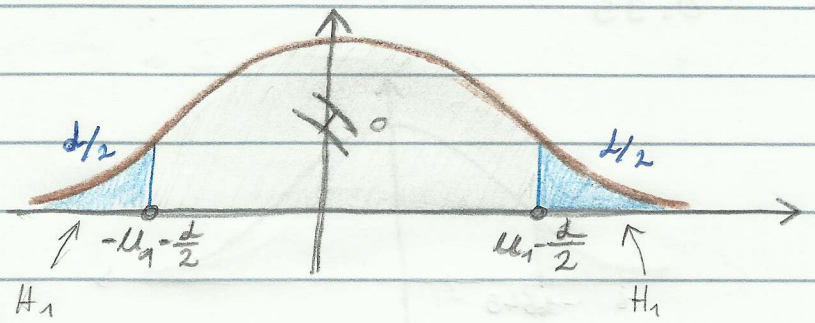
$H_0 \ldots \quad \vartheta = \vartheta_0$

$H_1 \ldots \quad \vartheta \neq \vartheta_0$ (dvostrani test)

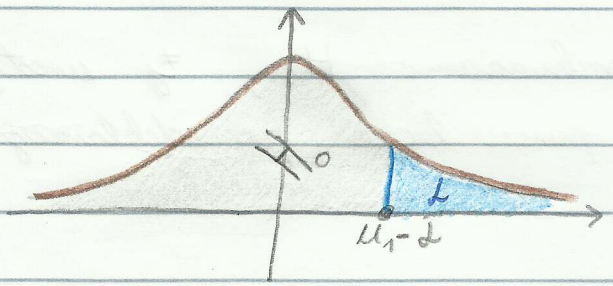
$\vartheta > \vartheta_0$ (desni test)

$\vartheta < \vartheta_0$ (lijevi test)

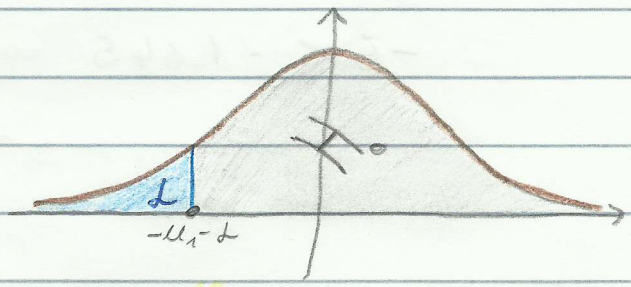
↳ nakon postavljanja ispravnog testa, treba izračunati statistiku iz uzorka ($U, T, \ldots$)



dvostrukI test



desni test



lijevi test

Zad: a) Utvrditi da punjenje štrcaljki boca radi sa devijacijom 5 ml. Uzorak od 100 boca dao je srednju vrijednost 997 ml. Provjeriti uz nivo značajnosti 5% da li ovaj uzorak govori u prilog pune boce.

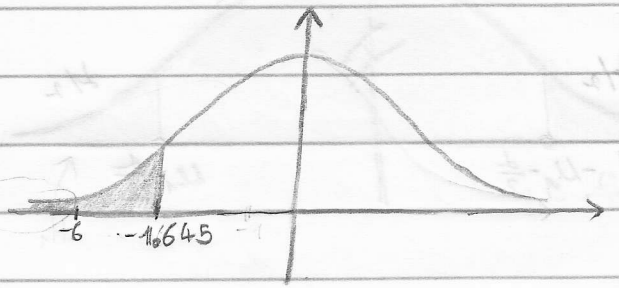
$H_0 \dots \quad a = 1000$

$H_1 \dots \quad a < 1000$

$$\hat{u} = \frac{\bar{x} - a_0}{\sigma/\sqrt{n}} = \frac{997 - 1000}{5/\sqrt{100}}$$

$$\hat{u} = -6$$

$$-u_{1-\alpha} = -u_{0.95} = -1.645$$



$-6 < -1.645 \Rightarrow$ prihvaćamo H_1, tj. uzorak ne govori u prilog pune boce po deklaraciji

b) Na drugom uređaju istog tipa na uzorku od 80 boca dao je vrednju vrijednost od 1001 ml. Provjerite hipotezu da oba uređaja pune boce na jednak način uz nivo značajnost 5 % !

$H_0 \dots \mu = \nu$

$H_1 \dots \mu \neq \nu$ → sada imamo dvostrani test

$$\sigma_z^2 = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}$$

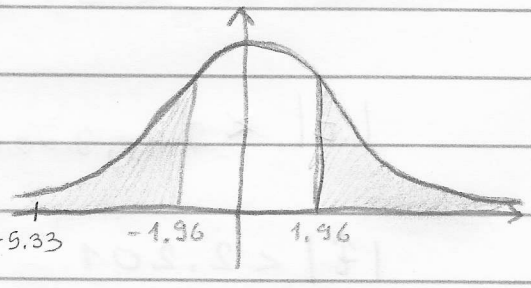
$n = 100$, $\sigma_x = 5$

$m = 80$, $\sigma_y = 5$

$$\sigma_z^2 = 0.5625$$

$$\hat{u} = \frac{\bar{X} - \bar{Y}}{\sigma_z} = \frac{997 - 1001}{\sqrt{0.5625}} = -5.33$$

$$u_{1-\frac{\alpha}{5}} = 1.96$$



$$|\hat{u}| > u_{1-\frac{\alpha}{5}}$$

$5.33 > 1.96$ → uz nivo značajnosti 5% možemo tvrdit da su uređaji različiti //

ZAD: Neka rafinerija proizvodi benzin za koji se očekuje oktanski broj 98. Uzeto je 12 uzoraka benzina i za njih su dobiveni sljedeći podaci: 99.1, 96.9, ..., 97.1, 97.6 (ovo nije niz, samo nisam zapisivao sve podatke ☺) Provjerimo ispravnost proizvodnje benzina! $\alpha = 0.05$

↳ postavimo hipoteze:

$H_0 \dots \quad a = 98$

$H_1 \dots \quad a \neq 98$

$$\hat{t} = \frac{\bar{X} - a}{\frac{S}{\sqrt{n}}}$$

$n = 12$

$\bar{X} = 97.6583$

$$\hat{S}^2 = \frac{1}{n-1}\left(\sum x_i^2 - n \cdot \bar{x}_i^2\right)$$

$\hat{S}^2 = 0.6079$

$\hat{t} = -1.518$

↳ za ovaj test, računamo da trebamo kvantil $t_{n-1,\,1-\frac{\alpha}{2}}$:

$t_{11;\,\alpha=0.05} = 2.201$

$|\hat{t}| < t_{11,\,\alpha=0.05}$

$|\hat{t}| < 2.201 \Rightarrow$ prihvaća se H_0, pa je proizvodnja ispravna

ZAD: 90 od 200 muškaraca i 69 od 150 žena je rekla da se bavi sportom. Da li možemo, uz značajnost 5%, zaključiti da se jednak postotak žena i muškaraca bavi sportom?

↳ ovo su dvije populacije i trebamo hipotezu o jednakosti proporcija:

- postavimo hipoteze:

$H_0 \dots p_1 = p_2$

$H_1 \dots p_1 \neq p_2$

$\hat{p}_1 = \frac{90}{200} = 0.45$ $\hat{p}_2 = \frac{69}{150} = 0.46$

↳ zajednička proporcija: $p = \frac{159}{350} = 0.454$

↳ disperzija: $\sigma^2 = p(1-p) = 0.248$

↳ trebamo i $\hat{u}$:

$$\hat{u} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} = -0.186$$

↳ imamo dvostran test, pa gledamo:

$$|\hat{u}| > u_{1-\frac{\alpha}{2}}$$

$|\hat{u}| > 1.96 \Rightarrow 0.186 > 1.96$

↓

ovo ne vrijedi, pa se prihvaća H_0 (jednak broj žena i muškaraca se bavi sportom)

ZAD: Proizvođač tvrdi da uređaj radi 200 dana. Ispitali smo 8 uređaja i dobili smo sljedeće podatke:

167, 171, 183, 187, 194, 204, 206, 240.

Uz razinu značajnosti 10%, je li proizvođač u pravu ili nas laže?

↳ ovdje trebamo hipotezu o očekivanju uz nepoznatu disperziju:

- postavimo hipoteze:

$H_0 \dots \; a = 200$

$H_1 \dots \; a < 200$

$n = 8$

$\bar{X} = 190.25$

$\hat{S}^2 = 259.36$

$$\hat{t} = \frac{\bar{X} - a}{\hat{S}/\sqrt{n}}$$

$$\hat{t} = -1.712$$

$$\hat{t} < -t_{n-1,\,1-\alpha}$$

$$-1.712 < -t_{7,\,1-\alpha}$$

$$-1.712 < -t_{7,\,1-0.1}$$ } PAZI!

↳ tablica je napravljena za $\alpha/2$, pa ćitaj stupac 0.2 //

$$-1.712 < -1.415$$

↑ ovo vrijedi, pa ne prihvaćamo H_1 //

ZAD: Od 97 studenata, 69 je položilo. Uz koj nivo značajnosti možemo tvrdit da predmet ima Bolonsku prolaznost od barem 75 % ?

↳ sada treba reći za koj „α" neka hipoteza vrijedi :-

· hipoteze: $H_0 \ldots p = 0.75$

$H_1 \ldots p < 0.75$

$$\hat{u} = \left(\frac{m}{n} - p_0\right)\sqrt{\frac{n}{p_0 q_0}} \quad ; \quad \hat{p} = \frac{m}{n} = 0.711$$

$$\hat{u} = (0.711 - 0.75) \cdot \sqrt{\frac{97}{0.75 \cdot 0.25}}$$

$$\hat{u} = -0.879$$

↳ moramo prihvatit H_0:

$$\hat{u} > -u_{1-\alpha}$$

$$-\hat{u} < u_{1-\alpha} \Rightarrow 0.879 < u_{1-\alpha}$$

↑ trebamo uzeti prvi broj veći od 0.879:

$$1 - \alpha = 0.82$$

$$\alpha = 0.18$$

↑ trebamo nivo značajnosti od 18%

Zad: Na 30 pacijenata, testiran je lijek X. Dobiveno je prosječno smanjenje od 23 jedinice tlaka uz devijaciju 3.2. Uz nivo značajnosti 5%, provjeri je li taj lijek bolji od starog lijeka koji smanjuje tlak za 21 jedinicu?

↳ testirana očekivanje - prosječno smanjenje krvnog tlaka ☺

• NAPOMENA: imamo samo jedan uzorak od 30 ljudi, pa koristi te formule ☺

↳ postavimo hipoteze:

$H_0 \dots \quad a = 21$

$H_1 \dots \quad a > 21$

↳ NAPOMENA: devijacija je neporemećena! Mi imamo samo devijaciju iz uzorka ☺ (stoga, koristimo t - test)

$n = 30 \Rightarrow \bar{X} = 23, \quad \hat{s} = 3.2$

$$\hat{t} = \frac{\bar{X} - a}{\hat{s}/\sqrt{n}} = 3.423$$

$$t_{29,\,1-0.05} = 1.699$$

$$\hat{t} > t_{29,\,1-0.05} \Rightarrow 3.423 > 1.699$$

↑ odbacuje se H_0, pa je novi lijek efikasniji

Zad 8 Imamo sljedeće podatke:

A	12	9	10	8	7	13	11	x
B	10	8	9	7	12	10	7	y

U dvije trgovine prodaje se čokolada marke P. Prodaja tijekom tjedna dana je tablicom. Može li se uz razinu značajnosti 10% zaključiti da se prosječna dnevna prodaja čokolade u ovim trgovinama razlikuje?

$H_0 \dots \mu = \nu$ $\quad n = 7, \; m = 7$

$H_1 \dots \mu \neq \nu$ $\quad \bar{X} = 10, \; \bar{Y} = 9$

$$\hat{S}_x^2 = \frac{1}{n-1}\left(\sum x_i^2 - n\bar{X}^2\right) = \frac{28}{6}; \quad \hat{S}_y^2 = \frac{20}{6}$$

$$\hat{S}_z^2 = \frac{1}{n+m-2}\left[(n-1)\hat{S}_x^2 + (m-1)\hat{S}_x^2\right] = 4$$

$$\hat{t} = \frac{\bar{X} - \bar{Y}}{\hat{S}_z^2} \cdot \sqrt{\frac{nm}{n+m}} = 0.935$$

$$|\hat{t}| > t_{n+m-2,\, 1-\frac{\alpha}{2}}$$

$$|\hat{t}| > t_{12,\, \alpha = 10\%}$$

$$0.935 > 1.782$$

↑

ovo ne vrijedi, pa se prihvaća hipoteza H_0, te je prodaja čokolade P u ta dva dućana jednaka

ZAD: Prilikom ispitivanja uzorka proizvoda, utvrđeno je da je svaki 4 proizvod škart. Proizvođač tvrdi da je škarta 20%. Koliko minimalno mora biti uzorak da bi uz značajnost 5% mogli zaključiti da proizvođač laže?

$H_0 \dots \quad p = 0.2$

$H_1 \dots \quad p > 0.2$

$\hat{p} = 25\% = 0.25$

$$\hat{u} = (\hat{p} - p) \cdot \sqrt{\frac{n}{p \cdot q}}$$

$$\hat{u} = 0.05 \cdot \sqrt{\frac{n}{0.2 \cdot 0.8}}$$

↳ mi trebamo odabrat H_1, pa mora vrijediti:

$$\hat{u} > u_{1-\alpha}$$

$$0.05 \cdot \sqrt{\frac{n}{0.2 \cdot 0.8}} > u_{0.95}$$

$$0.05 \cdot \sqrt{\frac{n}{0.2 \cdot 0.8}} > 1.64485$$

$$\frac{n}{0.2 \cdot 0.8} > 1082.2126$$

$$n > 173.15$$

$$n = 174$$

↑

moramo uzet 174 proizvoda

TEST PRILAGODBE RAZDIOBAMA

⇒ pitamo se:

- da li se neki podac stvarno ravnaju po nekoj zadanoj razdiobi?
  ↳ npr. po binomnoj, Poissonovoj, ...

ZAD: Bilježen je broj krivo spojenih poziva i dobiveni su sljedeći podac:

x_j	n_j	p_j
0	95	
1	140	
2	73	
3	31	
4	16	
5	7	
6	2	
7	0	
8	1	

broj dana

teorijske vjerojatnost

broj krivo spojenih poziva

Ravna li se razdioba po Poissonovoj razdiobi uz nivo značajnosti 0.02?

OKRENI →

↳ prvo trebamo izračunati teorijske vjerojatnosti:

• no, za to trebamo „λ" jer je:

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\lambda = \overline{X} = 1.364$$

x_j	n_j	p_j	np_j	$\frac{(n_j - np_j)^2}{np_j}$
0	95	0.2556	93.31	0.031
1	140	0.3486	127.25	1.277
2	73	0.2377	86.777	2.187
3	31	0.108	39.45	1.81
4	16	0.037	13.45	0.483
5	7	0.01	3.67	} 6.008
6	2	0.0023	0.83	
7	0	0.000445	0.163	
8	1	0.000076	0.028	
	≈ 1		≈ 365	

↳ spajamo sve razrede sa susjednima u kojima imamo manje od 5 elemenata sve dok nemamo 5 il više u razredu

• stoga, to je zadnjih 4 stupca

↳ spajanje se vrši tek nakon računanja „np_j" stupca, ali prije „$(n_j - np_j)^2 / np_j$" stupca

$\chi^2_g = 11.796$

$m = 6$, → broj redaka nakon spajanja

$r = 1$

↳ sada moramo nać vrijednost za ispitivanje hipotere:

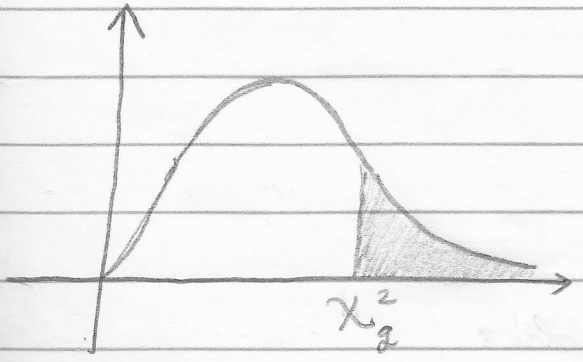
$$\chi^2_{m-r-1,\,1-\alpha} = \chi^2_{4,\,0.98} = 11.668$$

↳ odbacivanje hipotere

$$\chi^2_g > \chi^2_{m-r-1,\,1-\alpha}$$

$11.796 > 11.668$

↖ H_0 se odbacuje, pa se podaci ne ravnaju po Poissonovoj razdiobi



ZAD: Ponavljanjem pokusa, dobivene su sljedeći podaci varijable X:

$[a, b]$	0.-25	25-50	50-75	75-∞
n_j	87	46	14	3

Odredi, ravna li se X po eksponencijalnom razdiobi s nivoom značajnosti 5%!

↳ trebamo naći sredine danih intervala i izgradit tablicu kako bi to mogli upisat u kalkulator - to se radi jerlja jer je varijabla kontinuirana

- za radnji interval uzimamo jednaku duljinu kao i za ove prethodne, ako to već nije zadano

↳ gradimo tablicu:

$\overline{x_j}$	n_j	...
12.5	87	
37.5	46	
62.5	14	
87.5	3	
$\sum$	$n=150$	...

↳ za eksponencijalnu razdiobu:

$$f(x) = \lambda e^{-\lambda x}, \quad x>0$$

$$F(x) = 1 - e^{-\lambda x}$$

$$\lambda = \frac{1}{\overline{x}} = \frac{1}{26.33}$$

$$\lambda = 0.038$$

p_j	np_j
0.613	91.99
0.237	35.576
0.092	13.76
0.035	5.32
≈ 1	

$P(0 < X < 25) = 0.613$

$$P(a < X < b) = F(b) - F(a) = e^{-\lambda a} - e^{-\lambda b}$$

↳ spajanje razreda i računanje $\frac{(n_j - np_j)^2}{np_j}$:

$\frac{(n_j - np_j)^2}{np_j}$
0.271
3.054
0.227
$\chi^2_p = 3.552$

$m = 3$

$r = 1$

$$\chi_{m-r-1,\,1-0.05} = 3.841$$

↳ vidimo da je $\chi^2_p < \chi_{m-r-1,\,1-0.05}$, pa se H_0 ne odbacuje već se prihvaća, te se zaključuje da se podaci ravnaju po eksponencijalnoj razdiobi.

ZAD: 100 puta bacamo 3 kocke i broj šestica je dan ovako:

0	1	2	3
61	29	8	2

Provjeri hipotezu o ispravnosti ovih kocaka na nivo značajnosti 5%.

x_j	n_j	p_j	np_j
0	61	$\left(\frac{5}{6}\right)^3$	57.87
1	29	$\binom{3}{1}\cdot\frac{1}{6}\left(\frac{5}{6}\right)^2$	34.72
2	8	$\binom{3}{2}\left(\frac{1}{6}\right)^2\cdot\frac{5}{6}$	6.94
3	2	$\left(\frac{1}{6}\right)^3$	0.46
Σ	$n=100$	1	

↳ u ovdje radimo trebamo provjeriti da li je ovo poravnato po binomnoj razdiobi: $B\left(3, \frac{1}{6}\right)$

↳ spajamo razrede:

n_j	np_j	$\frac{(n_j - np_j)^2}{np_j}$
61	57.87	0.169
29	34.72	0.942
10	7.4	0.913
100	≈ 100	$\chi^2 = 2.024$

$m = 3$

$r = 0$

$\chi^2_{m-r-1,\,1-\alpha} = 5.991$

$$\chi^2 = 2.024 < \chi^2_{m-r-1,\,1-\alpha} = 5.991$$

↑ kocke su ispravne

ZAD: U prvih 800 decimala broja π, znamenke su se pojavile na sljedeći način:

0	1	2	3	4	5	6	7	8	9
74	92	83	79	80	73	77	75	76	91

Provjeri da li je statistički ur uvo značajno od 10% jednako očekivano pojavljivanje svake znamenke ?

↳ gradimo tablicu:

x_s	n_s	p_s	np_s	$\frac{(n_s - np_s)^2}{np_s}$
0	74	1/10	80	0.45
1	92	1/10	80	1.8
2	83	1/10	80	0.1125
3	79	1/10	80	0.0125
4	80	1/10	80	0
5	73	1/10	80	0.6125
6	77	1/10	80	0.1125
7	75	1/10	80	0.3125
8	76	1/10	80	0.2
9	91	1/10	80	1.5125

$m = 10$

$r = 0$

$\chi^2 = 5.125$

$\chi^2_{9,0.9} = 14.684$

$\chi^2 < \chi^2_{9,0.9}$

π je normalan !