

TARTU ÜLIKOO
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Armand Sutt
Covid andmete põhjal SIR mudeli koostamine
Matemaatiline statistika
Bakalaureusetöö (9 EAP)

Juhendaja: Sven Laur PhD

TARTU 2022

COVID ANDMETE PÕHJAL SIR MUDELI KOOSTAMINE

Bakalaureusetöö

Armand Sutt

Lühikokkuvõte

Covid-19 puhangust inspireerituna sooviti uurida SIR-mudeli (nakatumata, nakatunud, tervenenu) kasutamist Covid-19 ja teiste viirushaiguste levikul. Selleks loodi kõigepealt lihtsakoeline diskreetse ajaga mudel, mille alusel inimestele juhuslikult kindel kontakte määratakse ja nad jäävad fikseeritud tõenäosusega haigeks. Haigusperioodi jooksul inimesed võivad olla kõigepealt nakkusohutud ja seejärel muutuvad nad nakkusohtlikuks. Seejärel muudetakse nad tervenenuks. Pärast lihtsama mudeli koostamist taandati mudel kõigepealt neljale muutujale, et saaks suuremamahulisi populatsioone arvutis kergemini simuleerida. Sellest omakorda arendati välja mudelid, millega sai kirjeldada vanuse-/riskigruppide erinevuseid ja omavahelise suhtluse mõju nakkusele. Kirjeldatakse ka mudeleid, mis kasutavad pidevat aega. Töö viimane osa toob näiteid, kus mudelit rakendatakse.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad: SIR-mudel, pandeemia, logistiline kasv.

CREATING A SIR MODEL BASED ON COVID DATA

Bachelor thesis

Armand Sutt

Abstract

Inspired by the outbreak of Covid-19, the author of this thesis wished to describe the applications of the SIR-model (*Susceptible, Infected, Recovered*) to predict and simulate the developments of the coronavirus and others. To achieve this, at first the author created a naive model with discrete time, which maps each individual with a fixed number of contacts and in the case of a connection between a uninfected and sick individual assigns the susceptible individual as sick. During the period of illness, the inhabitants of the simulated world may be at first infected without being infectious, after which they become spreaders. After a certain number of days, each infected individual recovers and is no longer susceptible nor a spreader. The naive model is then reduced to four key

variables which are helpful when simulating a population of a larger size. Another use for the reduced model is that it allows the author to derive formulas which can be used to describe different age groups and the effects of their intermingling. A way of simulating the spread of disease with time as a continuous variable is then described by the author. Finally, there are examples of various applications for the model.

CERCS research specialisation: P160 Statistics, operations research, programming, financial and actuarial mathematics.

Key Words: SIR-model, pandemic, logistic-growth.

Sisukord

Sissejuhatus	4
SIR Mudeli algoritm	5
Tähistused ja mudeli üldine kirjeldus	5
Kontaktimaatriks ja nakatumine	7
Nakatamise tõenäosus	8
Nelja parameetriga mudel	11
Riskigruppidega mudel	17
Nelja parameetriga mudel riskigruppide jaoks	21
SIR Mudel diferentsiaalvõrrandina	24
Mudelite kasutamine	26
Saaremaa	26
Koolilapsed	32
Haiged loobuvad $\frac{1}{2}$ tõenäosusega külastuskäigust	33
Kokkuvõte	34
Lisa 1	35
Kasutatud allikad	37

Sissejuhatus

2020. aasta detsembris avastati esimene *Severe Acute Respiratory Syndrome Coronavirus-2* või tuntumalt Covid või koroonaviiruse juhtum Hiinas Wuhani piirkonnas. Sellest ajast alates on maailm ning pigem just läänemaailm leidnud olulisi puudujääke oma võimekusest tegeleda pandeemiaga. Nende hulka võivad kuuluda liiga liberaalne riigikord, ravimite või haiglakohtade puudus, aga nagu ka sõjas, on üks väga oluline edu kriteerium luureinfo. Viiruse "sõdureid" reaalselt jälgida on keeruline, sest neid ei näe palja silmaga ega satelliidilt, mistõttu tuleb haiguspuhanguid kohelda teisel viisil.

Nii vaktsiinide välja töötajate kui ka tervisehoiu strateegide kasutuses on suur kogus andmeid haigusjuhtumite kohta. See tähendab, et vastava luureinfo saamiseks tuleb kasutada andmeid mudelite loomiseks. Üks levinumaid mudeleid, mida laialdaselt tänapäeva meditsiinistatistikas viiruste leviku kirjeldamiseks kasutatakse on SIR-tüüpi epideemia mudel, mida tutvustatakse järgnevas töös.

SIR Mudeli algoritm

Järgnev alapeatükk on kirjutatud järgneva teose põhjal:

- Easley ja Kleinberg, 2010.

Sellest käsiraamatust saab töö oma üldise idee SIR-mudeli loomiseks, aga selles peatükis kasutatavad mudelid ja nende valemid on originaalsed. Mudelite täpsemat koodi näeb Sutt, 2022, githubi hoidlas.

Tähistused ja mudeli üldine kirjeldus

Selleks, et simuleerida viiruste levikut, tuleb jagada populatsioon, kus on N inimest kolme gruppi.

- $\mathcal{S}$ - nakatumata inimesed (inglise keeles *susceptible*).
- $\mathcal{X}$ - Nakatunud inimesed (inglise keeles *infected*).
- $\mathcal{R}$ - surnud/tervenenud ja nakkusohutud inimesed (inglise keeles *recovered/removed*).

Allikmaterjalides kasutatakse nakatunud inimeste hulga tähistamiseks tavaliselt $\mathcal{I}$ aga antud tekstis otsustasime kasutada $\mathcal{X}$, sest i on ka laialdaselt levinud järjekorranumber.

Epidemiade reaalsel kulgu modelleerides tuleb arvesse võtta ka neljandat gruppi $\mathcal{Q}$, mis tähistab inimesi, kes määratakse karantiini. Sellesse hulka võib kuuluda ükskõik, millisest eelnevast grupist indiviide. Naiivses mudelis seda gruppi ei käsitleta.

Grupp $\mathcal{X}$ jaotub omakorda kaheks grupiks:

- $\mathcal{X}^+ \subseteq \mathcal{X}$ - Nakatunud indiviidid, kes on nakkusohtlikud (inglise keeles *spreaders*).
- $\mathcal{X}^- \subseteq \mathcal{X}$ - Nakatunud indiviidid, kes on nakkusohutud (inglise keeles *non-spreaders*).

Eelnimetatud gruppide puhul kehtivad järgmised tingimused:

- $\mathcal{X}^+ \cup \mathcal{X}^- = \mathcal{X}$ - Nende hulkade kõik indiviidid teevad kokku hulga $\mathcal{X}$.
- $\mathcal{X}^+ \cap \mathcal{X}^- = \emptyset$ - Indiviid ei saa olla korraga nakkusohtlik ja nakkusohutu.

Indiviid saab korraga kuuluda ainult ühte hulkadest $\mathcal{S}, \mathcal{X}, \mathcal{R}$. Teisisõnu inimene on kas ainult nakatumata, ainult nakatunud või ainult paranenud.

Käesolevas töös käsitletakse indiviide asemel nende järjekorranumbreid (nõ nimesid) ja esitatakse kõik eelpoolmainitud hulgad järjekorranumbritena, see tähendab $\mathcal{S}, \mathcal{X}, \mathcal{R} \subseteq \mathbb{N}$. Seega näiteks i -s indiviid on nakatumata parajasti siis kui $i \in \mathcal{S}$.

Hulkade $\mathcal{S}, \mathcal{X}, \mathcal{R}$ suurusi ajal t , ehk indiviidide arve, kes kuuluvad vastavatesse gruppidesse hetkel t tähistatakse järgnevalt:

$$S(t) = |\mathcal{S}(t)|, \quad X(t) = |\mathcal{X}(t)|, \quad R(t) = |\mathcal{R}(t)|.$$

Kõikide gruppide suuruste summa on alati populatsiooni suurus N

$$S(t) + X(t) + R(t) = N.$$

Protsessi simuleerimiseks GNU R keskkonnas on kasulik tuua sisse staatusvektor. Kui maailmas on N indiviidi, siis vektor $\vec{w} = (w_1, w_2, \dots, w_N)$ on staatusvektor. Sellel staatusvektoril on neli võimalikku väärtust $w_i \in \{0, 1, 2, 3\}$. Siin 0 tähistab indiviidi, kes on nakatumata ehk $w_i = 0$ parajasti siis kui $i \in \mathcal{S}$. Tähisega 1 ja 2 tähistatakse vastavalt $w_i = 1 \Leftrightarrow i \in \mathcal{X}^-$ ja $w_i = 2 \Leftrightarrow i \in \mathcal{X}^+$. Tervenened/surnud indiviidide puhul kehtib $w_i = 3 \Leftrightarrow i \in \mathcal{R}$.

Mudelis, mida selles peatükis käsitletakse, rakendatakse aega diskreetselt, ehk teisisõnu maailm liigub edasi sammudena $t = 0, t = 1, t = 2, \dots$. Seega tegelikult näiteks $X = X(t)$, mis on nakatunute arv päeval t . Ühte sammu käsitletakse antud töö kontekstis ühe päevana. Kui inimene on olnud nakatunud kindla arvu samme T_1 , siis inimene terveneb või sureb ning liigub hulka $\mathcal{R}$. Ühtlasi, kui nakatunud inimene olnud nakatunud kindla arvu samme T_2 , kus $0 < T_2 \leq T_1$, siis nakatunud inimene muutub nakkusohtlikuks. Siin simuleeritakse olukorda, kus inimene on nakkusohutu või haigusel on peiteperiood. Näiteks COVID muutub väga nakkusohtlikuks 2 päeva enne sümptomeid, aga sümptomite esinemise aja mediaan on 5 – 6 päeva, seega saab taolisi nähtusi selles töös simuleerida.

Simulatsiooni alguses määratakse igale maailma nakatunud indiviidile nakatunud olemise päevade arv ajavektorina $\vec{T} = (t_1, t_2, \dots, t_N)$. Kui inimene i nakatub, siis tema aeg $t_i = T_1$ hakkab tiksuma, millelt igal päeval lahutatakse maha suurus 1. Kui on maha lahutatud väärtusi nii palju, et $T_1 - t_i = T_2$, siis muutub inimene nakkusohtlikuks $i \Rightarrow \mathcal{X}^+$. Kui inimese nakatumisest on

möödas T_1 päeva, siis ta muutub taas nakkusohutuks, aga teda ei saa ka enam nakatada, mis tähendab, et ta antud simulatsiooni kontekstis suundub tervenenu/surnute hulka $i \implies \mathcal{R}$

Inimeste omavahelise kohtumiste simuleerimiseks kasutatakse järgnevat lihtsat mudelit. Mudelis defineeritakse arv C , mis näitab, mitmele inimesele läheb keegi külla ühe päeva jooksul. Sündmust, kus indiviid i läheb külla indiviidile j tähistatakse $i \rightarrow j$. Kui eelnimetatud sündmus toimub ning üks neist on nakatumata ja teine nakatunud, siis kehtib tõenäosus ρ , et nakatumata indiviid nakatub.

Käesolev mudel ei ole 100% adekvaantne, sest siin ei ole tegelikult arvestatud inimeste reaalsete suhtlusvõrgustikega nagu on seda tehtud tekstis Easley ja Kleinberg (2010). Teisisõnu antud mudelis ei arvestata, et inimesed lävivad suurema tõenäosusega võrdlemisi staatilistes suhtusklastrites (sõbrad, perekonnad, töökaaslased) ja inimestele määratakse kontaktid juhuslikult. Teine oluline erinevus on, et siin on kõik inimesed võrdsed reaalsuses on erakuid ja supersuhtlejaid. Edasised mudelid hoiavad ühes grupis olevaid inimesi ikka samasugustena. Individuaalseid supersuhtlejaid ja erakuid siin töös ei käsitleta.

Kontaktimaatriks ja nakatumine

Olgu maailma staatusvektor $\vec{w} = (w_1, w_2, \dots, w_N)$. Kui on fikseeritud arv kontakte C , siis antud simulatsioonis tähendab, et iga indiviid i läheb iga päev külla täpselt C -le indiviidile. Ühtlasi võtab ta vastu kelle iganes, kes soovib talle külla tulla.

Selleks luuakse igal päeval $N \times C$ kontaktimaatriks K , kus igale reale i genereeritakse C arvu $k_{ij} \in \{1, \dots, N\}$, mis näitab, kellele vastava rea järjekorranumbriga indiviid külla läheb.

Kui panna kontaktid kirja maatriksina, siis see näeks välja selline:

$$K = \begin{pmatrix} k_{11} & \dots & k_{1C} \\ k_{21} & \dots & k_{2C} \\ \dots & \dots & \dots \\ k_{N1} & \dots & k_{NC} \end{pmatrix},$$

kus esimene veerg väljaspool sulge näitab, millisele indiviidile kontakt genereeritakse ning rida maatriksi kohal näitab, mitmes kontakt päevas see tema jaoks on. Igale reale genereeritakse C arvu ühtlasest jaotusest ning see näitab, kellele käesoleva päeva jooksul indiviid i külla läheb.

Igale lahtrile genereeritakse juhuslik arv eraldi ning see tähendab, et mitte ühegi indiviidi külaskäigud ei sõltu teiste külaskäikudest. Näiteks on imeväike tõenäosus, et indiviid i käib ühel päeval C korda külas indiviidil j ja indiviid j käib samal päeval C korda külas indiviidil i .

Kontaktimaatriksi genereerimiseks kasutatakse statistilise programmeerimiskeele R käsku `runif`, mis genereerib arve ühtlasest jaotusest vastavalt ülatoodud kirjeldusele:

```
generate_contacts <- function(n, c){
  contact_map <- rep(1:n, each=c)
  added_contacts <- round(runif(n*c)*n, 0)%n+1
  contact_list <- matrix(c(contact_map, added_contacts), ncol=2, byrow=F)
  return(contact_list)
}
```

Selline protsess annab ühtlase jaotuse järgi arve hulgast $\{1, \dots, N\}$ nagu mudelis vaja on.

Reaalses maailmas ei tohiks inimesele i genereerida vastavusse teda ennast, ehk sündmus $i \rightarrow i$ ei tohiks leida aset, sest inimene ei käi iseendal külas küll aga võib ühele inimesele võib minna rohkem kui üks kord külla. Antud programmis valik, mis lubab seda, sest selline implementatsioon on mugavam ja seda on hiljem lihtsam arvutustes käsitleda. Teisisõnu kui vaadata sündmust $i \rightarrow j$, siis reaalses maailmas oleks $P(i \rightarrow j) = \frac{1}{N-1}$, aga antud juhul $P(i \rightarrow j) = \frac{1}{N}$. Kuna suurte arvude puhul $N \approx N - 1$, siis on protsessid väga sarnased. Valim luuakse tagasipanekuga.

Nakatamise tõenäosus

Olgu $\mathcal{X}(t)$ ja $\mathcal{X}(t+1)$ nakatunute hulgad ajahetkedel t ja $t+1$. Lühiduse mõttes tähistagu $i \implies \mathcal{X}(t+1)$ sündmust, et indiviid i nakatub ajahetkel t haigusesse. Formaalselt $i \implies \mathcal{X}(t+1)$ parajasti siis kui $i \in \mathcal{X}(t+1) \setminus \mathcal{X}(t)$. Kui ajahetk $t+1$ on kontekstist teada, siis saab kasutada veel lühemat tähistust $i \implies \mathcal{X}$.

Kui vaadelda nakatumata indiviidi i , siis mis peaks juhtuma, et ta nakatuks? Selleks, et indiviid i nakatuks, peab ta sattuma kontakti teiste indiviididega, kes on nakkusohtlikud ning seejärel nakatuma. Olgu ρ vastav haiguse ülekandumise tõenäosus. See väärtus ei pruugi olla iga inimese puhul, aga antud mudeli puhul tehakse lihtsustus, kus määratakse kõigile üks “keskmine” nakatumistõenäosus. Meditsiinilist vaatepunktis saab nakatumist modeleerida kahel erineval moel. Kõige lihtsam on eeldada, et mitme nakkusohtliku inimesega kokku saades toimub nakkumine iga kohtumise ajal teistest kohtumistest sõltumatult. Bioloogiliselt pole see mudel korrektne,

kuna erinevatelt inimestelt saadud viirusdoosid kumuleeruvad ja 10 nakatunuga kohtumise korral peaks nakatumistõenäosus olema vähemalt 10 korda kõrgem kui ühe nakatunuga kohtudes. Seda saab mudeldada liites kokku iga kohtumise nakatumistõenäosused. Rangelt võttes on vastav nakatumistõenäosus suurem kui nakatumiste sõltumatust eeldades.

Sõltumatu nakatumismudel. Paneme tähele, et indiviid i saab nakatuda kahel teineteist mitte välistaval moel: minnes külla nakatunule (sündmus A) või võõrustades nakatunud külalist (sündmus B).

Leiame esmalt tõenäosuse, et terve inimene nakatub k -ndal külaskäigul. Kuna nakkusohtlikke inimesi on X^+ , siis on vastav tõenäosus

$$p_1 = \frac{X^+}{N} \cdot \rho$$

Kuna nakatumised on sõltumatud teineteisest, siis tõenäosus, et indiviid i ei nakatunud ühegi külaskäigu ajal on

$$p_2 = \left(1 - \frac{X^+}{N} \cdot \rho\right)^C$$

Järelikult tõenäosus, et i läheb C kontakti puhul külla mõnele nakkusohtlikule indiviidile on

$$\Pr[A] = 1 - \left(1 - \frac{X^+}{N} \cdot \rho\right)^C$$

Teine võimalus nakatuda on läbi külaliste võõrustamise. Tõenäosus, et kõikide inimeste k -nda visiidi ajal tuleb indiviidil i võõrustada suvalist haiget indiviidi j on $\frac{1}{N}$. Siin aga tuleks tähele panna, et nakatumata indiviidile i võivad juhuslikult külla sattuda ka kõik nakkusohtlikud indiviidid ja vastavad sündmused pole teineteist välistavad. Seetõttu on kasulik arvutada tõenäosus, et indiviid j ei suuda teda k -nda visiidi ajal nakatada:

$$q_1 = 1 - \frac{\rho}{N}$$

Nüüd saab sündmuste sõltumatust ära kasutades lihtsalt leida, et peale kõikide külaliste võõrustamist on i ikka veel nakatumata

$$q_2 = \left(1 - \frac{\rho}{N}\right)^{X^+ \cdot C}$$

Järelikult tõenäosus, et indiviid i nakatus läbi võõrustamise

$$\Pr[B] = 1 - \left(1 - \frac{\rho}{N}\right)^{X^+ \cdot C}.$$

Kuna igale indiviidile genereeritakse kontaktid juhuslikult, siis sündmused A ja B on teineteisest sõltumatud. seega tõenäosus, et indiviid i nakatub on

$$\Pr[i \implies \mathcal{X}] = \Pr[A] + \Pr[B] - \Pr[A] \cdot \Pr[B]. \quad (1)$$

Kumulatiivne nakatumismudel. Kumulatiivse nakatumismudeli korral tuleb meil kõik individuaalsete nakatumissündmuste tõenäosused omavahel kokku liita ja skaleerida parameetriga ρ .

Osasündmuste summa, kus indiviid i nakatub, sest ta läks külla mõnele indiviidile, kes oli nakkusohtlik on

$$p_A = \frac{X^+ \cdot C}{N}$$

Osasündmuste summa, kus indiviid i nakatub, kuna ta võõrustas mõnda nakkusohtlikku indiviidi on

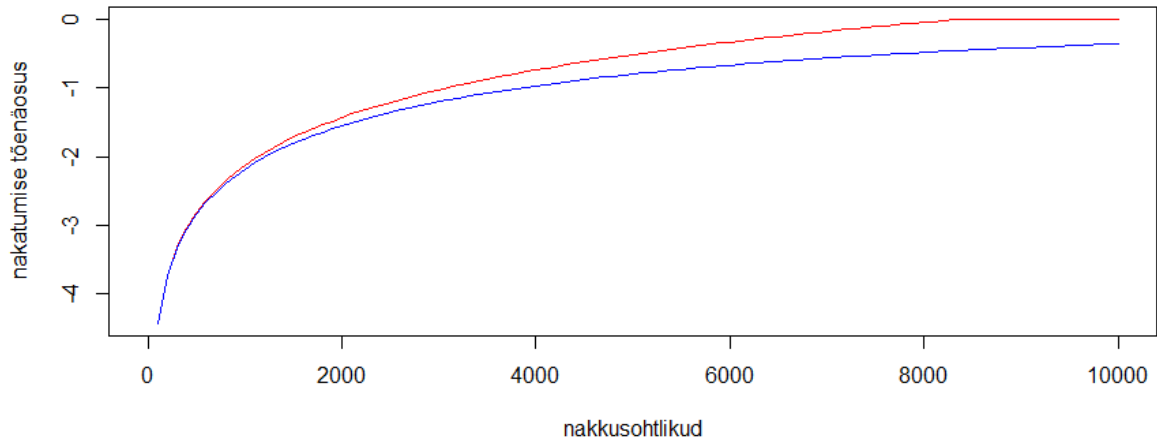
$$p_B = \frac{X^+ \cdot C}{N}$$

Kuna võib juhtuda, et $\rho \cdot C \cdot X^+ > N$, siis selleks, et nakatumise tõenäosus ei läheks suuremaks kui 1 kasutatakse valemina järgmist valemit:

$$\Pr[i \implies \mathcal{X}] = \min\{\rho \cdot (p_A + p_B), 1\} \quad (2)$$

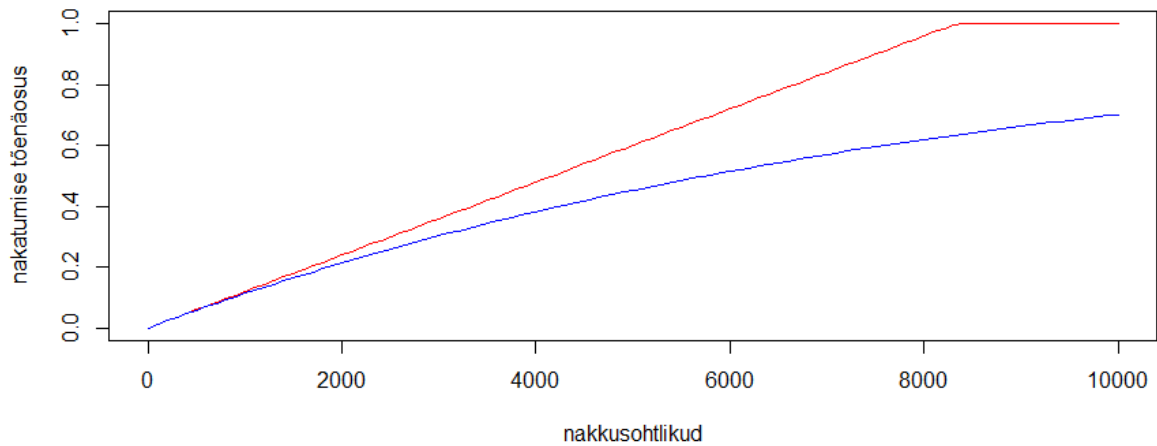
Siit on näha, et sõltumatu ja kumulatiivse mudeli erinevus hakkab rolli mängima suuremate X^+ väärtuste juures. Kui nakkusohtlikuid on $\sim 10\%$, siis hakkavad tõenäosused oluliselt erinema. Seda seetõttu, et kumulatiivne mudel on sisuliselt lineaarne, aga sõltumatu mudel seda pole. See on oluline, sest kumulatiivne mudel jõuab mingilt maalt 1-ni, aga sõltumatu mudel mitte.

??



Joonis 1: Mudelite võrdlus $N = 100000$, $C = 20$, $\rho = 0.3$, tõenäosus logaritmitud e baasil.

??



Joonis 2: Mudelite võrdlus $N = 100000$, $C = 20$, $\rho = 0.3$, tõenäosus logaritmitamata e baasil.

Nelja parameetriga mudel

Olenemata nakatumismudelitest on igal tervel indiviidil on võrdne tõenäosus nakatuda ning see tõenäosus sõltub ainult suurusest X^+ . Seega on igal sammul lisanduvate nakatuvate inimeste arv binoomjaotusega sündmus. Selle jaotuse esimene parameeter on $S(t)$ ja sõltumatu mudeli puhul teine parameeter on $p(t) = \Pr[i \Rightarrow \mathcal{X}(t)]$.

Kõige lihtsamal juhul, kui inimene on nakkusohtlik vaid järgmisel päeval peale nakkust, saab taandada maailma kolmele muutujale $S(t)$, $X(t)$, $R(t)$. Üldisel juhul on tarvis teada ka igapäevast uute nakatunute arvu $\Delta X(t)$.

Teisisõnu kui $X(t)$ on nakatunute arv ajahetkel t , siis selleks, et saada uued nakatunud ajahetkel t , tuleb genereerida arv $\Delta X(t)$, on võimalik genereerida juhuslik arv binoomjaotusest:

$$\Delta X(t) \sim \text{Binom}(n = S(t), Pr = p(t))$$

Paneme nüüd tähele, et suuruse $\Delta X(t)$ läbi saab leida ka teiste parameetrite väärtused. Kuna kõik individid hulgas $\mathcal{X}$ tulevad hulgast $\mathcal{S}$, siis

$$S(t+1) = S(t) - \Delta X(t) \ .$$

$X(t)$ puhul tuleks meele pidada, et igal ajahetkel on $X(t+1)$ summa $X(t)$ -st, $\Delta X(t)$ -st, aga samas võetakse ka maha tervenenu inimesed $\Delta R(t)$. Seega

$$X(t+1) = X(t) + \Delta X(t) - \Delta R(t) \ .$$

Kuna meie mudeli järgi saab inimene terveks T_1 päeva pärast nakatumist, siis

$$\Delta R(t) = -\Delta X(t - T_1)$$

ning seetõttu

$$R(t+1) = R(t) + \Delta R(t) = R(t) + \Delta X(t+1 - T_1) \ .$$

Seega saab süsteemi arengut kirjeldada järgneva võrduste blokiga

$$S(t+1) = S(t) - \Delta X(t) \ ,$$

$$X(t+1) = X(t) + \Delta X(t) - \Delta R(t) \ ,$$

$$R(t+1) = R(t) + \Delta X(t+1 - T_1)$$

.

Selleks et igal ajahetkel leida nakatumistõenäosust $p(t)$ peame veel lisaks eelnevale teadma $X^+(t)$ väärtust.

$$X^+(t+1) = X^+ + \Delta X^+(t) = X^+ + \Delta X(t+1-T_2) - \Delta X(t-T_2)$$

Teades nakkusohtlike arvu, on võimalik välja arvutada $p(t)$, mis avalduvad valemites (4) ja (2).

Sõltumatu mudel Antud mudelis pannakse arvutused samamoodi kokku nagu eelnevas peatükis arutletud. Kõigepealt arvutatakse $P_A = \Pr[A]$ ja seejärel $P_B = \Pr[B]$ ning siis pannakse need vastavalt kokku.

```
p_ind <- function(s,x,xpos,r,c,rho){ #xpos = nakkusohtlike inimeste arv
  n = sum(c(s,x,xpos,r))
  a = (xpos/n)*rho #tõenäosus, et indiviid i nakatub minnes külla
      #mõnele nakkusohtlikule indiviidile
  a1 = 1-a
  a1c = exp(k*(log(1-a)))
  a1c1 = 1-a1c #tõenäosus, et nakkus leiab aset vähemalt ühel korral C-st
      #võimalikust kontaktist

  P_A = a1c1

  b = rho*1/n #tõenäosus, et nakkusohtlik mõni
      #indiviid j läheb külla ja nakatab indiviidi i
  b1 = 1 - b
  b1xd = eksp(xpos*log(1-b))
  b1xd1 = 1-b1xd
  b1xd11 = 1-b1xd1
  b1xd11c = b1xd11^c
  b1xd11c1 = 1-b1xd11c #tõenäosus, et ükskõik, milline nakkusohtlik indiviid läheb külla
      #ja nakatab indiviidi i

  P_B = b1xd11c1
  nak = P_A + P_B - P_A*P_B

  return(nak)
```

```
}
```

Seejärel selleks, et arvutada haigestunud inimeste arv, ehk siis $\Delta X(t)$ kasutatakse

```
n_of_inf2 <- function(s,x,xd,r,c,rho){  
  P = p_of_inf2(s,x,xd,r,c,rho)  
  deltaX = rbinom(1,s,P)  
  return(nak)  
},
```

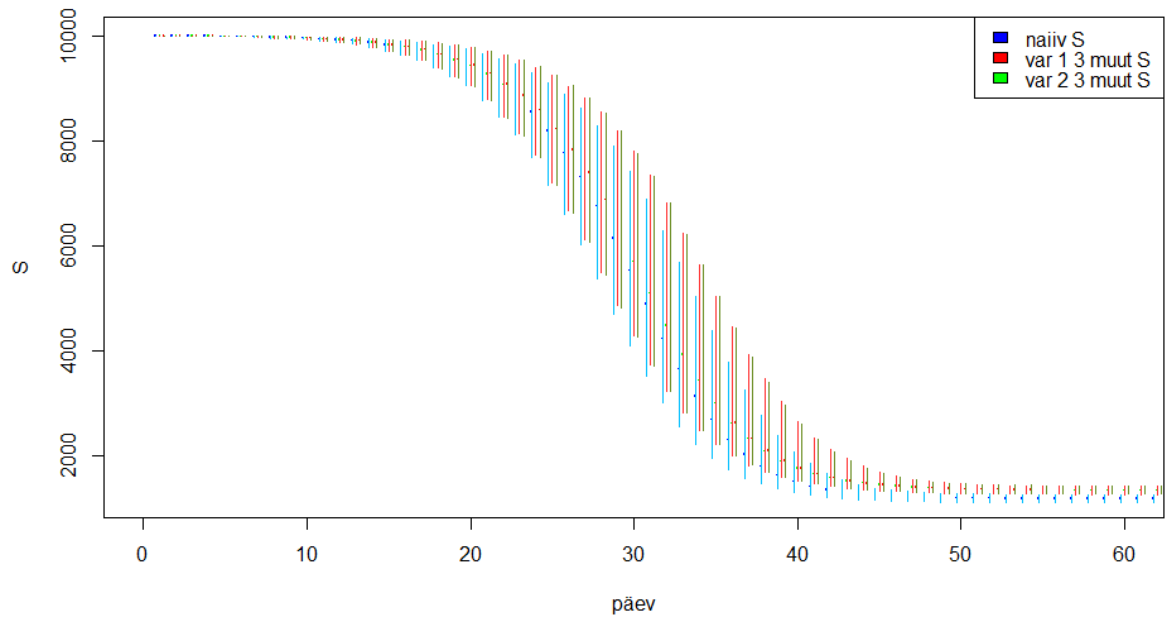
kus eelnevas koodijupis `rbinom` on programmi R käsk, mis genereerib juhuslikke arve binoomjao-
tusest.

Kumulatiivne mudel Kumulatiivse mudeli puhul tehakse arvutused kumulatiivse nakatu-
mistõenäosuse valemi järgi, kus $pa = \Pr[A]$ ja $pb = \Pr[B]$:

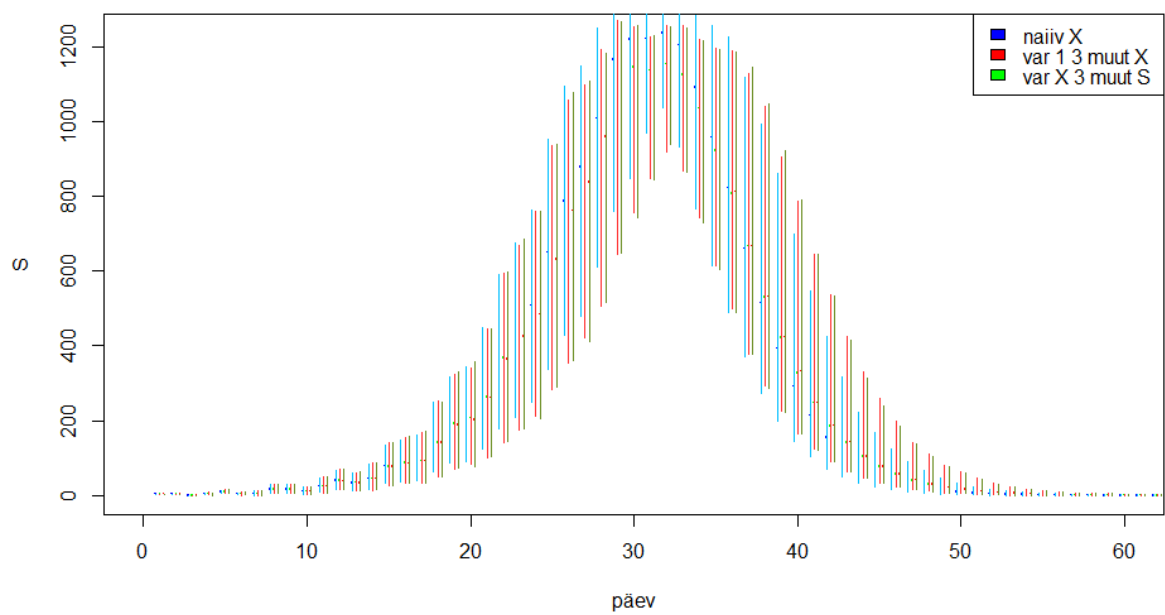
```
p_cum<-function(s,x,pos,r,c,rho){  
  n <- sum(c(s,xd,x,r))  
  pa <- pos*c/n  
  pb <- pos*c/n  
  papb <- pa+pb  
  pnak <- min(c(rho*papb,1))  
  return(rbinom(1,s,pnak))  
},
```

kus arvutatakse eelkirjeldatud viisil välja mõlemad tõenäosused P_A ja P_B ning R-i käsuga `rbinom`
genereeritakse iga sammu jaoks uute nakatunute arv.

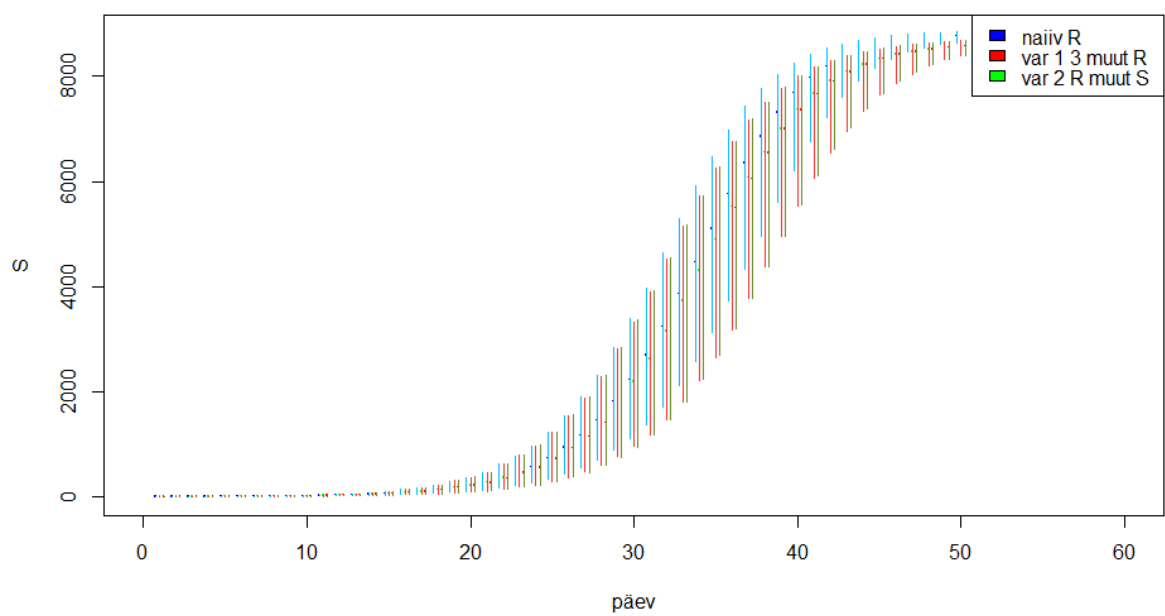
Järgmistel joonistel on tehtud iga meetodiga 1000 simulatsiooni ja on kuvatud nende mediaantu-
lemus ning 5% ja 95% kvantiilide vahemikud.



Joonis 3: Nakatumata inimeste arv S



Joonis 4: Nakatunud inimeste arv X



Joonis 5: Tervenenuid inimeste arv R

Nagu näha, siis naiivne mudel nakatab täpselt nende parameetritega populatsiooni kiiremini, kui teised mudelid.

Selleks, et T_1 ja T_2 juhuslikult varieerida, võib igal sammul võtta nendeks mõne juhusliku arvu ja siis lisada saadud arv vastavasse järjendisse τ_1 ja τ_2 . Seejärel tuleb igal sammul lisatud nakatunude arv $\Delta X(t)$ lisada teise järjendisse. Igal sammul lahutatakse järjenditest τ_1 ja τ_2 maha 1 ning kui mõni τ_2 jõuab nulli, siis lahutatakse sellele positsioonile vastav $\Delta X(t)$ maha nakkusohtlike seast $X^-(t)$ ja kui mõni τ_1 jõuab nulli, lahutatakse sellele vastav väärtus maha $X^+(t)$ seast. Seda aga selles töös ei tehta.

Riskigruppidega mudel

Tavaliselt on maailmas erinevad grupid inimesi, kellel on erinev tõenäosus haigestuda või teisi nakatada. Näiteks võib jagada gruppidesse inimesi vanuse või teiste komplikatsioonide olemasolu järgi. Olgu vastavad riskigrupid $\mathcal{G}_1, \dots, \mathcal{G}_\ell$, kus loomulikult $\mathcal{G}_u \cap \mathcal{G}_v = \emptyset$ kui $u \neq v$ ning $\mathcal{G}_1 \cup \dots \cup \mathcal{G}_\ell$ katab kogu populatsiooni. Olgu iga baasgrupp $\mathcal{S}, \mathcal{X}, \mathcal{R}$ jaotatud omakorda alamgruppideks vastavalt riskigruppidele:

$$\mathcal{S}_u = \mathcal{G}_u \cap \mathcal{S} ,$$

$$\mathcal{X}_u = \mathcal{G}_u \cap \mathcal{X} ,$$

$$\mathcal{R}_u = \mathcal{G}_u \cap \mathcal{R} .$$

Konstruksioonist lähtuvalt

$$\mathcal{S}_u \cap \mathcal{X}_u = \emptyset ,$$

$$\mathcal{X}_u \cap \mathcal{R}_u = \emptyset ,$$

$$\mathcal{S}_u \cap \mathcal{R}_u = \emptyset ,$$

ning

$$\mathcal{G}_u = \mathcal{S}_u \cup \mathcal{X}_u \cup \mathcal{R}_u .$$

Ühtlasi $S_u + X_u + R_u = N_u$.

Kui meid huvitab tõenäosus, et indiviid i , kes kuulub riskigruppi $\mathcal{G}_u$ nakatub, siis ta peab kokku saama kellegiga, kes on nakkusohtlik ja seega kuuluma ühte riskigruppi $\mathcal{X}_1^+, \dots, \mathcal{X}_\ell^+$.

Kuna aga erinevad riskigrupid lävivad üksteisega erineva tõenäosusega (näiteks nooremad inimesed käivad omavahel rohkem läbi, kui vanemate inimestega jne), siis seda tuleb seda tõenäosuslikult mudeldada. Selleks defineeritakse $\ell \times \ell$ maatriks α_{uv} , kus u tähistab indiviidi grupist $\mathcal{G}_u$, kes läheb külla indiviidile grupist $\mathcal{G}_v$

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1\ell-1} & \alpha_{1\ell} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2\ell-1} & \alpha_{2\ell} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{\ell 1} & \alpha_{\ell 2} & \dots & \alpha_{\ell\ell-1} & \alpha_{\ell\ell} \end{pmatrix}.$$

Tuleb tähele panna, et tõenäosus, et indiviid i grupist $\mathcal{G}_u$ läheb külla indiviidile j grupist $\mathcal{G}_v$ ei sõltu indiviidi i ega j järjekorranumbritest. Pärast grupi valimist kaaluga α_{uv} külastatakse kõiki indiviide grupis $\mathcal{G}_v$ sama tõenäosusega.

Olgu meil teada, et indiviid $i \in \mathcal{G}_u$ külastas k -külaskäigul indiviidi grupist $\mathcal{G}_v$, siis tõenäosus, et vastav indiviid oli nakkusohtlik, on

$$\frac{X_v^+}{N_v}.$$

Seega on tõenäosus, et indiviid $i \in \mathcal{G}_u$ külastas k -külaskäigul indiviidi kes oli nakkusohtlik, on

$$\alpha_{uv} \cdot \frac{X_v^+}{N_v}.$$

Kui nakatumistõenäosus kahe kahe inimese vahel ei sõltu nende riskigruppidesse kuuluvusest, siis saame vastavaks nakatmistõenäosuseks

$$p_1 = \alpha_{uv} \cdot \frac{X_v^+}{N_v} \cdot \rho$$

Kuigi päris maailmas on igal indiviidil sisuliselt kindel viirusaine kogus, mille täitudes inimene nakatub ja see ei sõltu sellest, kas ta läheb kellegile külla või talle tuleb keegi külla. Selle töö kontekstis kasutatakse siiski viirusaine kogust ρ_{uv} , mis tähistab viirusainet olukorras, kus grupi u indiviid läheb külla grupi v indiviidile, sest sellega saab sisse arvestada tingimused, kus mõnes külastuse suunas pannakse ette mask, hoitakse distantssi või vastupidiselt ollakse just riskantsema käitumisega.

Tähistagu nüüd maatriks

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1\ell} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2\ell} \\ \dots & \dots & \dots & \dots \\ \rho_{\ell 1} & \rho_{\ell 2} & \dots & \rho_{\ell \ell} \end{pmatrix},$$

kus ρ_{uv} tähistab nakatumistõenäosust u läheb külla v kes on nakatunud ja nakatub.

Tähistagu

$$\vec{C} = (C_1, \dots, C_\ell)$$

mitu külastust teeb indiviid vastavast grupist. Seda on vaja, sest erinevad grupid võivad käituda erinevalt. Näiteks ühe klišeena võib tuua mõnele vanusegrupile omistatud motto "Äh, ma olen spordi poiss raudse tervisega, pole see Covid midagi..."

Sõltumatu nakkusmudel. Nagu varemgi vaatame esmalt olukorda kus erinevate nakkus-
ohtlike inimestega kohtumised ei mõjuta teineteist.

...

Arvestades eelnevat on tõenäosus, et inimene grupist $\mathcal{G}_u$ läheb külastusel k külla inimesele grupist $\mathcal{G}_v$ ja et too on nakkusohtlik on:

$$\alpha_{uv} \cdot \frac{X_v^+}{N_v}$$

ning seega tõenäosus, et ta nakatub külaskäigul k :

$$p_1 = \sum_{v=1}^{\ell} \alpha_{uv} \cdot \frac{X_v^+}{N_v} \cdot \rho_{uv}$$

Tõenäosus, et inimene grupist $\mathcal{G}_u$ ei nakatu sammul t kui ta teeb C_u külastust on:

$$p_2 = \left(1 - \sum_{v=1}^{\ell} \alpha_{uv} \cdot \rho_{uv} \cdot \frac{X_v^+}{N_v} \right)^{C_u}$$

Seega tõenäosus, et kui inimene grupist $\mathcal{G}_u$ teeb C_u külastust ja nakatub on:

$$\Pr[A|i \in \mathcal{G}_u] = 1 - \left(1 - \sum_{v=1}^{\ell} \alpha_{uv} \cdot \rho_{uv} \cdot \frac{X_v^+}{N_v}\right)^{C_u} \quad (3)$$

Kuna indiviidid grupist $\mathcal{G}_v$ võivad väga paljudel teineteist mittevälistavatel viisidel külastada indiviidi i , on kasulik arvutada tõenäosus, et nakkusohtlik indiviid $j \in \mathcal{X}_v^+$ ei suuda teda k -nda visiidi ajal nakatada:

$$q_{vu} = \begin{cases} 1 - \alpha_{vu} \cdot \frac{1}{N_u} \cdot \rho_{vu}, & \text{kui } k \leq C_v, \\ 1, & \text{kui } k > C_v. \end{cases}$$

Nüüd saab sündmuste sõltumatust ära kasutades lihtsalt leida, et peale kõikide külaliste võõrustamist grupist $\mathcal{G}_v$ on i ikka veel nakatumata:

$$\left(1 - \alpha_{vu} \cdot \frac{1}{N_u} \cdot \rho_{vu}\right)^{X_v^+}$$

Tuleb ka silmas pidada, et igal indiviidil riskigrupist $\mathcal{X}_v^+$ on võimalik kontakteeruda teistega C_v korda iga päev. Seega tõenäosus et ükski indiviid grupist $\mathcal{G}_v$ ei nakata inimest i terve päeva jooksul on

$$q_v = \left(1 - \alpha_{vu} \cdot \frac{1}{N_u} \cdot \rho_{vu}\right)^{X_v^+ C_v}$$

Järelikult tõenäosus, et indiviid i nakatus läbi võõrustamise, on

$$\Pr[B|i \in \mathcal{G}_u] = 1 - \prod_{v=1}^{\ell} \left(1 - \alpha_{vu} \frac{1}{N_u} \rho_{vu}\right)^{X_v^+ C_v}.$$

Kuna igale indiviidile genereeritakse kontaktid juhuslikult, siis sündmused A ja B on teineteisest sõltumatud ja seega

$$\Pr[i \implies \mathcal{X}|i \in \mathcal{G}_u] = \Pr[A|i \in \mathcal{G}_u] + \Pr[B|i \in \mathcal{G}_u] - \Pr[A|i \in \mathcal{G}_u] \cdot \Pr[B|i \in \mathcal{G}_u]. \quad (4)$$

Kumulatiivne nakkumismudel. Kumulatiivse mudeli puhul võetakse järgnevast kahest võrratusest paremad pooled ja liidetakse need kokku. Tõenäosus, et indiviid i nakatub, kui tal on C_u kontakti grupiga $\mathcal{G}_v$:

$$\Pr[A|i \in \mathcal{G}_u] \leq \sum_{v=1}^{\ell} C_u \frac{\alpha_{uv} X_v^+}{N_v} \rho_{uv} \ .$$

Tõenäosus, et indiviid i nakatub, kui temal käib külas grupi $\mathcal{G}_v$ indiviid on:

$$\Pr[B|i \in \mathcal{G}_u] \leq \sum_{v=1}^{\ell} C_v \frac{\alpha_{vu} X_u^+}{N_u} \rho_{vu} \ .$$

Tõenäosus, et indiviid i grupist $\mathcal{G}_u$ nakatub on

$$\Pr[i \implies \mathcal{X}|i \in \mathcal{G}_u] = \min\{sum_{v=1}^{\ell} C_u \frac{\alpha_{uv} X_v^+}{N_v} \rho_{uv} + \sum_{v=1}^{\ell} C_v \frac{\alpha_{vu} X_u^+}{N_u} \rho_{vu}, 1\}$$

Nelja parameetriga mudel riskigruppide jaoks

Sõltumata nakatumismudelitest, on igal indiviidil $i \in \mathcal{G}_u$ võrdne tõenäosus nakatuda ning see oleneb ainult eelnimetatud parameetritest. Seega igal sammul lisanduvate nakatuvate inimeste arv on binoomjaotusega sündmus. Selle jaotuse esimese parameeter on $S_u(t)$ ja sõltumatu mudeli puhul on teine parameeter

Kõige lihtsamal juhul, kui inimene on nakkusohtlik vaid järgmisel päeval peale nakkust, saab grupi $\mathcal{G}_u$ muutumise taandada maailma kolmele muutujale $S_u(t)$, $X_u(t)$, $R_u(t)$. Üldisel juhul on tarvis teada ka igapäevast uute nakatunute arvu $\Delta X_u(t)$.

Teisisõnu kui $X_u(t)$ on grupi $\mathcal{G}_u$ nakatunute arv ajahetkel t , siis selleks, et saada uued nakatunud ajahetkel $t+1$, tuleb genereerida arv $\Delta X_u(t)$, mida saab teha kasutades binoomjaotust:

$$\Delta X_u(t) \sim \text{Binom}(n_u = S_u(t), \text{Pr} = p(t))$$

Paneme nüüd tähele, et suuruse $\Delta X_u(t)$ läbi saab ka leida teiste parameetrite väärtused. Kuna kõik indiviidid hulgast $\mathcal{X}_u$ tulevad hulgast $\mathcal{S}_u$, siis

$$S_u(t+1) = S_u(t) - \Delta X_u(t)$$

$X_u(t)$ puhul tuleks ka silmas pidada, et igal ajahetkel on $X_u(t+1)$ summa $X_u(t)$ -st, $\Delta X_u(t)$ -st, aga ka tuleb maha lahutada need indiviidid, kes antud sammul tervenevad $\Delta R_u(t)$. Seega

$$X_u(t+1) = X(t) + \Delta X_u(t) - \Delta R_u(t)$$

Kuna meie mudeli järgi saab inimene terveks T_1 päeva pärast nakatumist, siis

$$\Delta R_u = \Delta X_u(t - T_1) ,$$

seega

$$R_u(t+1) = R_u(t) + \Delta X_u(t+1 - T_1)$$

Niiet seega saab grupi arengut kirjeldada järngevalt:

$$\begin{aligned} S_u(t+1) &= S_u(t) - \Delta X_u(t) , \\ X_u(t+1) &= X_u(t) + \Delta X_u(t) , \\ R_u(t+1) &= R_u(t) + \Delta X_u(t+1 - T_1) \end{aligned} \tag{5}$$

Kuna on vaja teada ka $X_u^+(t)$ väärtust, tuleb arvestada, et

$$X_u^+(t+1) = X_u^+(t) + \Delta X_u^+(t) = X_u^+ + \Delta X_u(t+1 - T_2) - \Delta X_u(t - T_2)$$

Mudeli arvutamine Selleks, et mudeli uusi väärtuseid leida, kasutatakse nii kumulatiivse kui ka sõltumatu mudeli jaoks sama viisi, kuidas korrigeerida gruppide suuruseid. Kõigepealt leitakse huvipakkuvad kordajad α maatriksist. Seejärel fikseeritakse külla mineja ja külla tuleja ρ -d ning seejärel skaleeritakse grupid. Suurus **gnr** on u , ehk grupp, kuhu kuulub indiviid i , kelle nakatumist uuritakse.

Sõltumatu mudel Selleks, et arvutada sõltumatu nakkusmudeli järgi nakatunude arv, kasutatakse igal sammul arvutamiseks järgnevat koodi:

```
nakatunude_arv <- function(svg,xvg,xdvg,rv,ck,gnr,alfa,rho){
```

```

probinfgnr <- rep(NA, times=length(svg))
probinfj <- 1
for(j in 1:length(svg)){
  probinfgnr[j] <- alfa[,gnr][j]*rho[,gnr][j]*xdvg[j]*(1/(sum(c(svg[j],xvg[j],xdvg[j],rvg[j]))))
  probinfj <- probinfj*((1-alfa[,j][gnr]*rho[,j][gnr]*(1/(sum(c(svg[gnr],xvg[gnr],xdvg[gnr],rvg[gnr]))))
}
p1 <- (1-sum(probinfgnr))*ck[gnr]
PrB <- 1-probinfj
PrA <- 1-p1
PrInf <- min(c(PrA+PrB-PrA*PrB,1))
return(round(rbinom(1,svg[gnr],PrInf)))
}

```

Kumulatiivne mudel Selleks, et arvutada kumulatiivse nakkusmudeli järgi nakatunude arv, kasutatakse järgnevat koodi:

```

nakatunude_arv2 <- function(svg,xvg,xdvg,rvg,ck,gnr,alfa,rho){
  PrA <- 0
  PrB <- 0
  for(j in 1:length(svg)){
    PrA = PrA + ck[gnr]*alfa[,gnr][j]*xdvg[j]*rho[,gnr][j]*1/(sum(c(svg[j],xvg[j],xdvg[j],rvg[j]))))
    PrB = PrB + ck[j]*alfa[,j][gnr]*xdvg[gnr]*rho[,j][gnr]*1/(sum(c(svg[gnr],xvg[gnr],xdvg[gnr],rvg[gnr]))))
  }
  PrInf <- min(c(PrA+PrB,1))
  return(rbinom(1,svg[gnr],PrInf))
}

```

SIR Mudel diferentsiaalvõrrandina

Järgnev peatükk on kirjutatud järgneva teose põhjal:

- Newman, 2010.

Varasema arutelu põhjal on teada, et nakatumise puhul on tegu binoomjaotusele vastava sündmusega. Eelnenud sisu üheks miinuseks on asjaolu, et aega käsitletakse diskreetsena, ehk samm-sammu haaval. Reaalsuses aga on aeg pidev, mis tähendab, et tegelikult võib toimuda nakatumine igal ajahetkel. Selleks, et nakatumist uurida kasutades pidevat aega, on mõistlik kasutada kumulatiivset mudelit, kuna pideva ajaga mudeli loomiseks tuleb luua diferentsiaalvõrrand ja sellele vastavat diferentsiaalvõrrandit on kõige lihtsam lahendada.

Kumulatiivses mudelis kasutatakse antud töö kontekstis uue nakatumise tõenäosuse arvutamiseks valemit:

$$Pr[i \in \mathcal{S} \Rightarrow \mathcal{X}] = \min \left\{ \frac{CX\rho}{n} + \frac{CX\rho}{n}, 1 \right\}$$

Selleks, et arvutada uute nakatunude arv, kasutatakse antud peatükis binoomjaotuse keskmist oodatavat tulemust niiet kui on $X_{nak} \sim Bin(S, Pr)$, siis kasutatakse väärtust

$$\begin{aligned} \frac{dX}{dt} &= \Delta X(t) - \Delta R(t) \\ &= S \cdot \min \left\{ \frac{CX\rho}{n} + \frac{CX\rho}{n}, 1 \right\} - \frac{dR}{dt} , \end{aligned} \tag{6}$$

et arvutada järgneva sammu haigestunud inimeste juurdekasv.

Tekstis Newman M.E.J. (2010) kasutatakse selleks valemit

$$\frac{dX}{dt} = \beta SX - \gamma X \tag{7}$$

Selleks, et antud töö valemit ?? samamoodi rakendada, tuleks see ühildada valemiga ?. β on

inimeste arv, kellega iga indiviid igal ajahetkel t kohtub, ehk antud töö C . Käesolevas töös kasutatakse ka väärtust ρ , mis on tõenäosus, et kontakti puhul indiviid nakkub, seega kasutatakse $\beta = C \cdot \rho$. See on kõige olulisem seos, mis on käesoleval tööil antud allikaga.

Antud tekstis on γ mingi osa nakatunudest, kes igal sammul paranevad. Meie puhul on väärtus T_1 , mis näitab, kui kaua on üks indiviid haige alates nakatumisest. Kui meil on indiviid $i \in \mathcal{X}$, kes peab paranema T_1 sammu jooksul, siis see tähendab, et iga päev paraneb indiviid i keskmiselt $\frac{1}{T_1}$ võrra. Teistpidi võib ka vaadata, et iga päev on tõenäosus $\frac{1}{T_1}$, et indiviid i paraneb. Seega eemaldunud indiviidide kasv $\Delta R \sim B(X, \frac{1}{T_1})$ ja võib öelda, et keskmiselt paraneb igal sammul $\frac{1}{T_1} \cdot X$ indiviidi. Seega antud juhul $\gamma = \frac{1}{T_1}$.

Lahendades need võrrandid Wolfram-Alphas saab:

$$S(t) = \frac{C_1}{e^{-C \cdot \rho \cdot t \cdot X(t)}} \quad (8)$$

$$X(t) = C_2 \cdot e^{C \cdot \rho \cdot t \cdot S(t) - \frac{1}{T_1} X(t)} \quad (9)$$

$$R(t) = C_3 + \frac{1}{T_1} \cdot t \cdot X(t) \quad (10)$$

Mudelite kasutamine

Antud peatükis kasutatakse loodud mudeleid, et vaadata kolme olukorda:

1. Simuleeritakse Saaremaal toimunud olukorda ja vaadatakse kuivõrd sõltub haiguse levik juhusest. Arvutatakse 90% kvantiilid ja keskmised päevade tulemused. Andmed on saadud Statistikaamet, 2022a, RV021: Saaremaa, vanuserühmad 0-85+.
2. Vaadeldakse olukorda, kus koolilapsed suhtlevad omavahel rohkem ja uuritakse, kuidas see mõjub vanavanematele. Andmed on saadud Statistikaamet, 2022b, RV021: Kogu eesti, vanuserühmad 0-85+.
3. Vaadeldakse olukorda, kus haiged inimesed loobuvad $\frac{1}{2}$ tõenäosusega külastuskäigust. Andmed on saadud Statistikaamet, 2022b, RV021: Kogu eesti, vanuserühmad 0-85+.

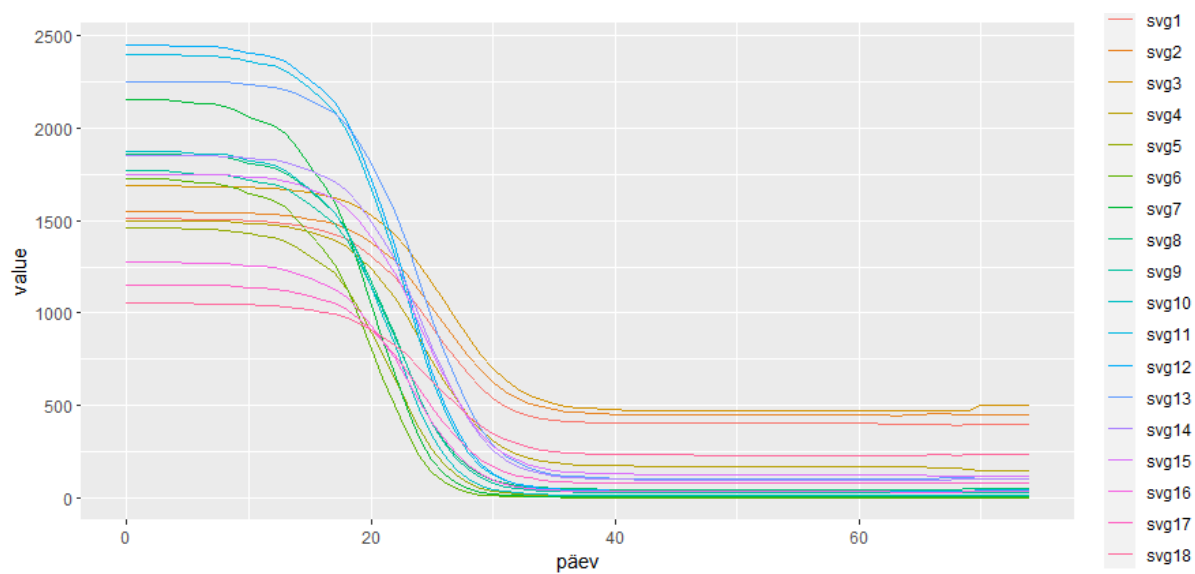
Saaremaa

Käesolevate andmete jaoks kasutatakse Statistikaameti andmeid tabelist RV0221U. Lisas 1 on lisatud ka nakkusmaatriks ρ , kohtumismaatriks α . Külastuste arvud igal vanusegrupil on $C = (5, 5, 5, 5, 20, 20, 15, 3, 5, 5, 3, 3, 2, 2, 2, 2, 2, 2)$.

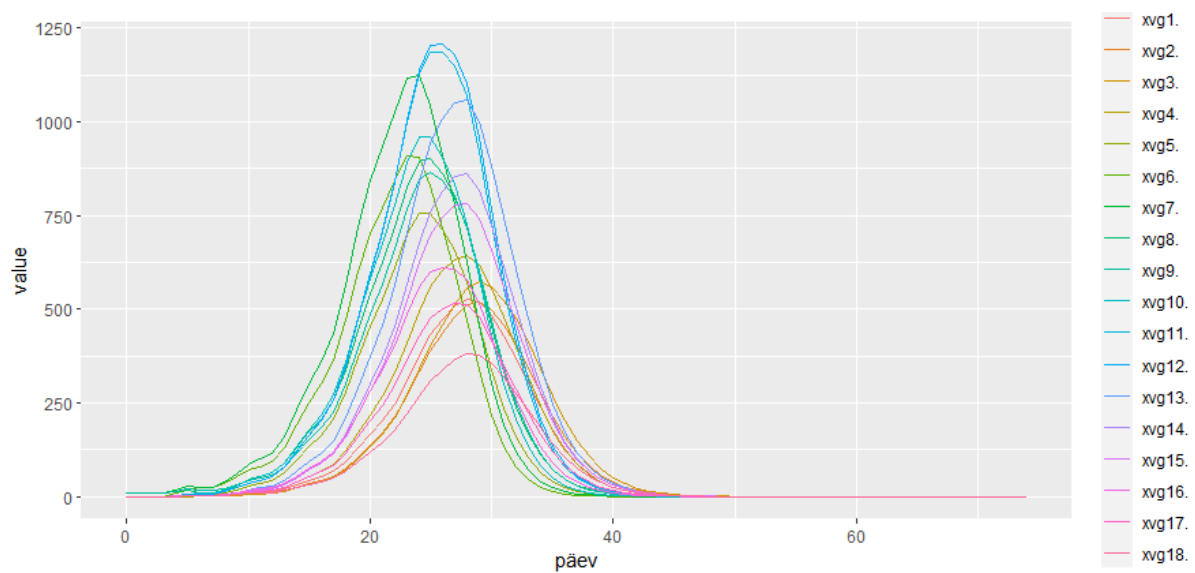
Vanuserühm	Muutuja	Elanikke
0-4	svg1/xvg1/rvg1	1498
5-9	svg2/xvg2/rvg2	1542
10-14	svg3/xvg3/rvg3	1682
15-19	svg4/xvg4/rvg4	1492
20-24	svg5/xvg5/rvg5	1437
25-29	svg6/xvg6/rvg6	1684
30-34	svg7/xvg7/rvg7	2123
35-39	svg8/xvg8/rvg8	1819
40-44	svg9/xvg9/rvg9	1719
45-49	svg10/xvg10/rvg10	1824
50-54	svg11/xvg11/rvg11	2341
55-59	svg12/xvg12/rvg12	2384
60-64	svg13/xvg13/rvg13	2199
65-69	svg14/xvg14/rvg14	1778
70-74	svg15/xvg15/rvg15	1669
75-79	svg16/xvg16/rvg16	1234
80-84	svg17/xvg17/rvg17	1118
85+	svg18/xvg18/rvg18	1031

Maatriks α sai valitud selline, et inimesed suhtlevad suurema tõenäosusega enda vanusegrupiga ja nende sarnastega ning oma vanemate põlvkonnaga. Ühtlasi suheldakse ka oma lastega ja lastelastega (Lisa 5).

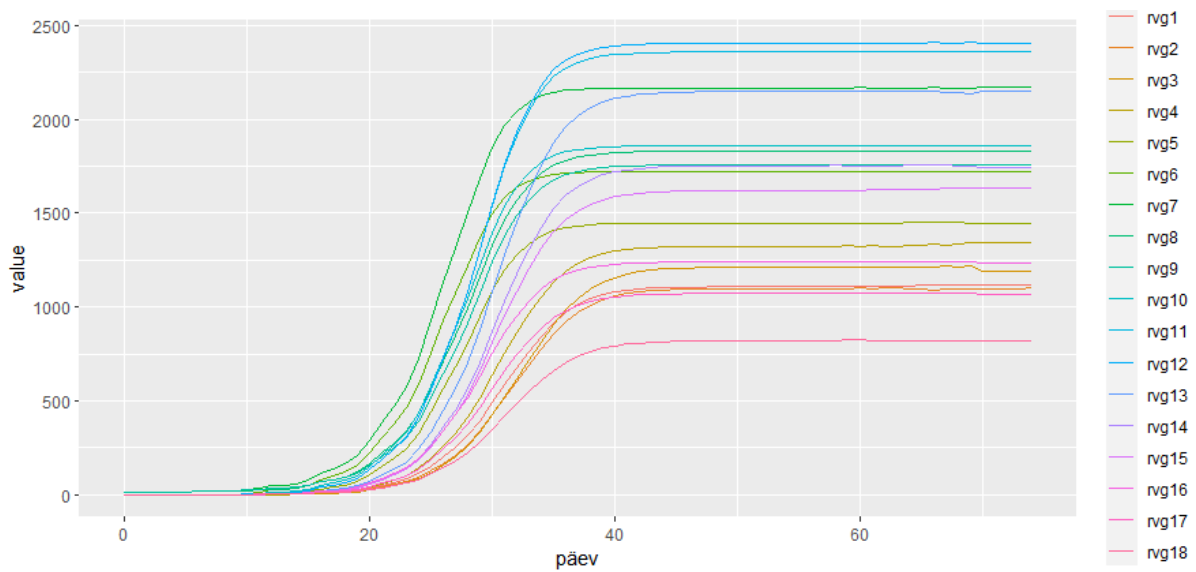
Vanusegruppide tulemused sõltumatu mudeliga Siin on kuvatud vanusegruppide keskmised tulemused päevade lõikes:



Joonis 6: Nakatumata inimeste arv S

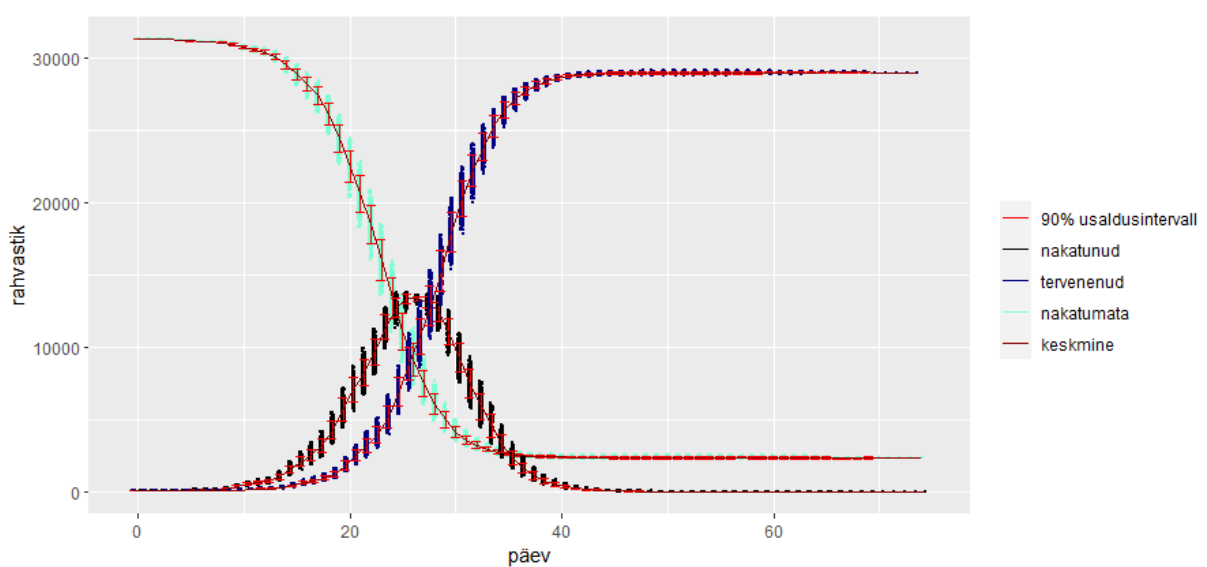


Joonis 7: Nakatunud inimeste arv X



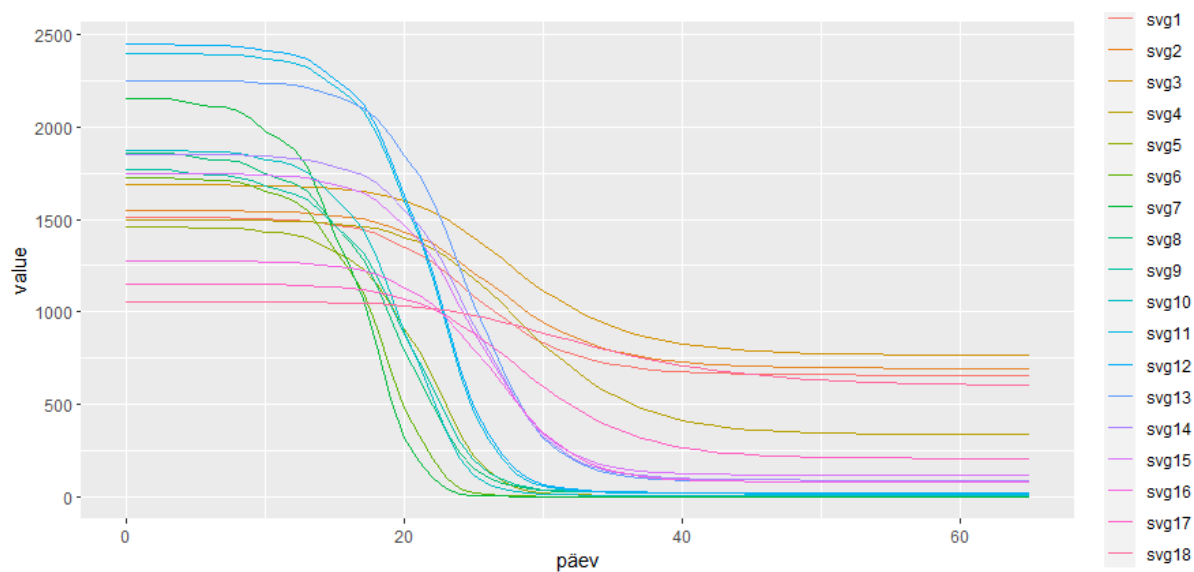
Joonis 8: Tervenenu inimeste arv R

Nakatunumata, nakatunud, tervenenu päevade lõikes koos kõikide simulatsioonide tulemustega, 90% usaldusintervallidega ning keskmise joonega:

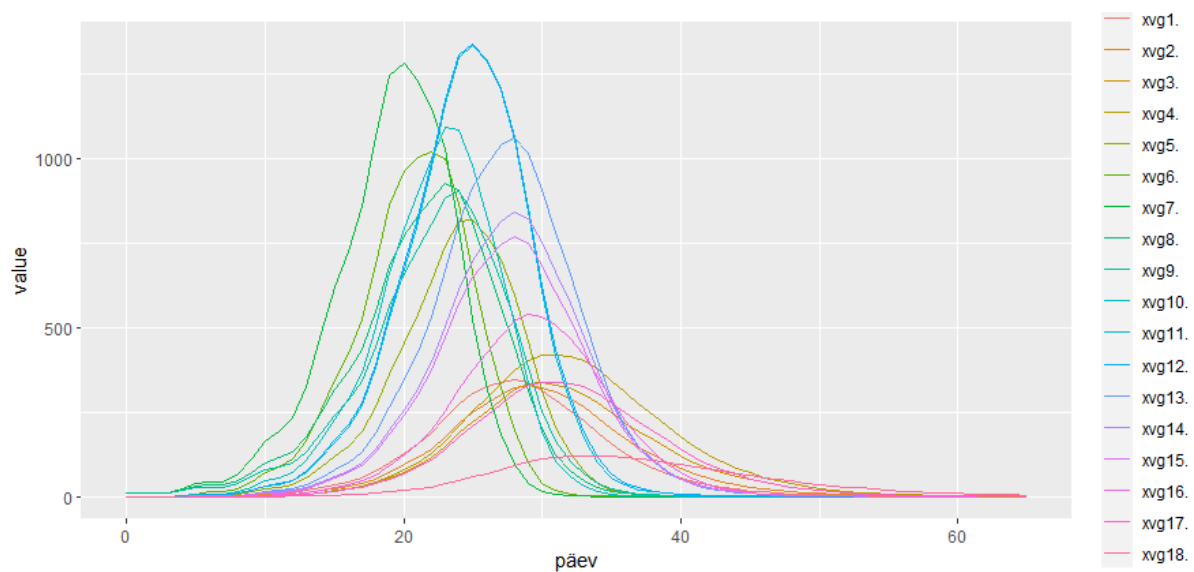


Joonis 9: Nakatunumata, nakatunude ning tervenenu keskmise ning 90% usaldusintervall

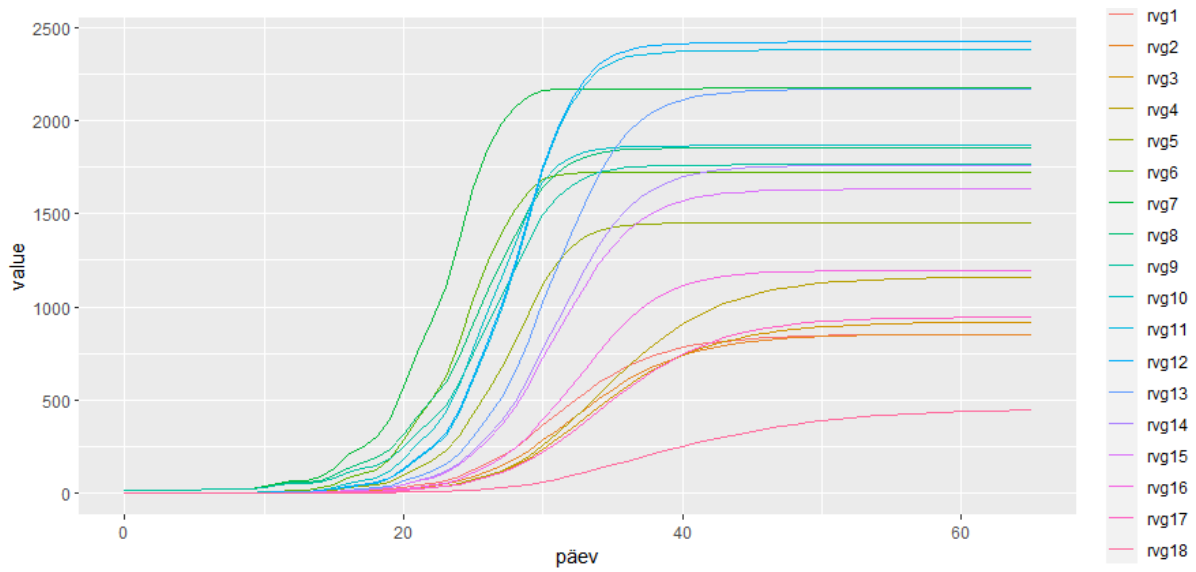
Vanusegruppide tulemused kumulatiivse mudeliga Siin on kuvatud vanusegruppide keskmised tulemused päevade lõikes:



Joonis 10: Nakatumata inimeste arv S

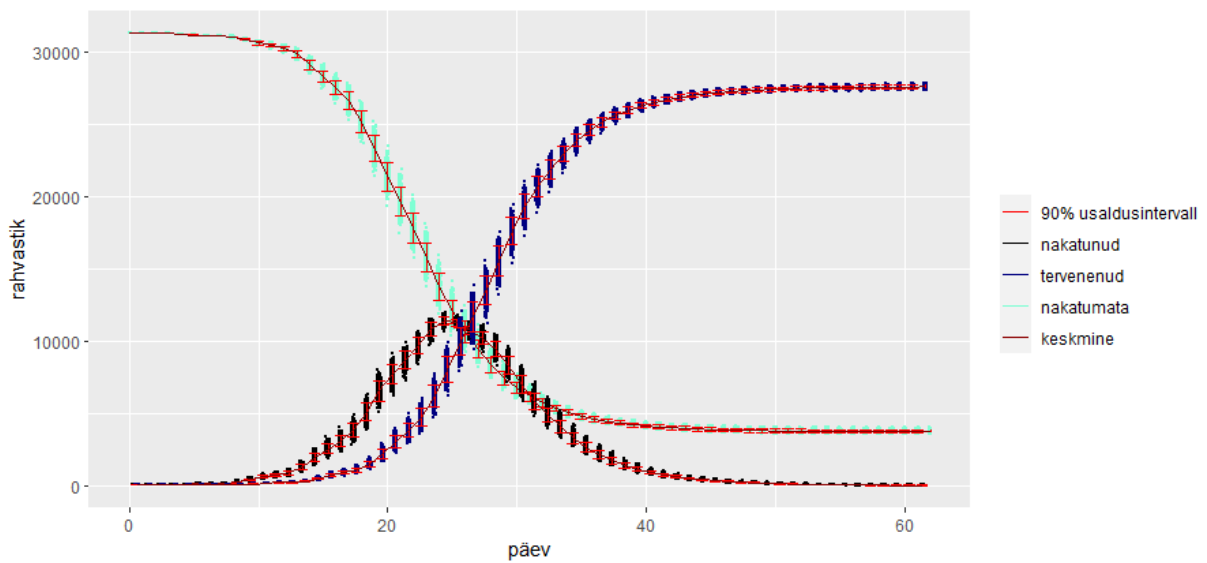


Joonis 11: Nakatunud inimeste arv X



Joonis 12: Tervenenu inimeste arv R

Nakatunumata, nakatunud, tervenenu päevade lõikes koos kõikide simulatsioonide tulemustega, 90% usaldusintervallidega ning keskmise joonega:



Joonis 13: Nakatunumata, nakatunude ning tervenenu keskmine ning 90% usaldusintervall

Meetodite statistilised näitajad

päev	$EX_{kumulatiivne}$	$\sigma_{kumulatiivne}$	90%	$EX_{sõltumatu}$	$\sigma_{sõltumatu}$	90%
0	30	0	30...30	30	0	30...30
7	159	12	139...179	140	12	120...159
14	1915	172	1633...2197	1549	160	1286...1813
21	8651	373	8038...9264	8193	537	7310...9076
28	9437	328	8897...9978	12427	393	11780...13073
35	2660	189	2349...2971	2376	305	1875...2878
42	606	52	521...692	151	28	106...196
49	153	19	122...184	8	3	3...13
56	42	9	27...57	1	1	0...2
63	12	5	4...20	1	< 1	0...2
70	4	3	0...8	...	...	...
77	2	2	0...5	...	...	...
84	2	1	0...4	...	...	...

Kumulatiivse ja sõltumatu mudeli võrdlemisel tuleb välja, et sõltumatu mudel jõuab kiiremini lõpuni kui kumulatiivne mudel. Teisisõnu kumulatiivne mudel on optimistlikum hinnang viiruse levikule ja sõltumatu mudel on pessimistlikum. Kumulatiivse mudeli puhul haigestus keskmiselt 340 inimest päevas ja sõltumatu mudeli puhul 510.

Koolilapsed

Mis saab siis, kui koolilapsed vanuses 5...9 kuni 15...19 suhtlevad 10 korda rohkem?

päev	$EX_{\text{norm. suhtlus}}$	$\sigma_{\text{norm. suhtlus}}$	90%	$EX_{\text{suuren. suhtlus}}$	$\sigma_{\text{suuren. suhtlus}}$	90%
0	30	0	30...30	30	0	30...30
7	90	10	74...105	157	12	138...176
14	278	35	221...335	2150	209	1806...2493
21	1473	208	1130...1816	18760	1766	15855...21666
28	6431	920	4917...7945	159713	13690	137193...182232
35	20241	2842	15565...24917	619802	12763	598807...640797

Nagu näha, siis suurendatud suhtluse puhul on viiruse levik väga laastav. 35. päevaks on

nakatunud umbes 620000 inimest võrreldes 20000 haigega kui inividid suhtleksid vähem.

Haiged loobuvad $\frac{1}{2}$ tõenäosusega külastuskäigust

Kui vanusegruppide kohtumistevektor on $C_{ei\ loobu} = (5, 5, 5, 5, 20, 20, 15, 3, 5, 5, 3, 3, 2, 2, 2, 2, 2, 2)$ selline, siis kui vähendada kõigi väärtuseid täpselt $\frac{1}{2} * C$ peale:

päev	$EX_{ei\ loobu}$	$\sigma_{ei\ loobu}$	90%	EX_{loobub}	σ_{loobub}	90%
0	30	0	30...30	30	0	30...30
7	133	11	115...152	67	8	53...81
14	1574	178	1281...1867	268	45	194...345
21	12125	1335	9929...14322	799	142	566...1032
28	93003	9723	77008...108998	2391	440	1667...3115
35	434839	23131	396789...472889	20906	3707	5050...9420

Antud tabelist on selgelt näha, et 35. viiruse leviku päeval on nakatunude erinevus 20 kordne nii, et need, kes ei jäta kohtumisi ära, saavad palju kõrgema haiguste voolu kui need, kes jätavad. Seega karantiini efekt on antud mudeli puhul selgelt näha.

Kokkuvõte

Käesoleva töö mudelitega saab võrdlemisi edukalt mudeldada viiruste käitumist populatsioonis, sest saab mudeldada vanusegruppe ja nendevahelist suhtlust ja gruppide karantiini. Vaevalt, et nende mudelite järgi saab riiklikku Covid-19 kaitset kavandada. Küll aga saab uurida üldiste kaitsemeetmete efektiivsust ja mõju populatsioonis ning on kasulik, et tutvustada põhilisi momente viiruse levikul ning nende mudeldamisel.

Sellegipoolest on antud töö puuduseid. Näiteks pole siin eristatud, kas inividid reaalselt surevad või tervenevad, seega jääb puudu taasnakatumise efekt. Välja on ka jäetud individuaalsed supersuhtlejad ja erakud ning ei mudeldata spetsiifilisi suhtlusvõrgustikke. See jätab mõneti reaalelulised tahud uurimata. Eelnimetatu aga polegi selle töö eesmärk.

Lisa 1

	[.1]	[.2]	[.3]	[.4]	[.5]	[.6]	[.7]	[.8]	[.9]	[.10]	[.11]	[.12]	[.13]	[.14]	[.15]	[.16]	[.17]	[.18]
[1,]	0.096	0.020	0.014	0.011	0.014	0.013	0.08	0.090	0.011	0.01	0.062	0.060	0.004	0.004	0.005	0.000	0.000	0.000
[2,]	0.029	0.204	0.058	0.045	0.014	0.013	0.01	0.011	0.085	0.08	0.008	0.007	0.059	0.062	0.004	0.010	0.000	0.000
[3,]	0.010	0.008	0.288	0.112	0.055	0.013	0.01	0.045	0.011	0.01	0.031	0.060	0.089	0.047	0.061	0.021	0.000	0.000
[4,]	0.024	0.008	0.006	0.225	0.137	0.050	0.01	0.011	0.011	0.01	0.155	0.075	0.030	0.031	0.020	0.021	0.016	0.012
[5,]	0.000	0.020	0.014	0.045	0.274	0.125	0.04	0.011	0.011	0.04	0.031	0.007	0.007	0.078	0.204	0.103	0.062	0.048
[6,]	0.000	0.004	0.029	0.045	0.014	0.125	0.20	0.112	0.106	0.10	0.062	0.060	0.030	0.031	0.020	0.021	0.016	0.000
[7,]	0.192	0.082	0.014	0.011	0.014	0.125	0.20	0.225	0.106	0.10	0.062	0.060	0.044	0.047	0.102	0.082	0.031	0.024
[8,]	0.192	0.163	0.058	0.045	0.137	0.250	0.10	0.112	0.085	0.08	0.031	0.030	0.030	0.031	0.020	0.021	0.093	0.024
[9,]	0.048	0.041	0.086	0.067	0.027	0.025	0.06	0.067	0.213	0.10	0.062	0.045	0.030	0.031	0.020	0.041	0.093	0.024
[10,]	0.024	0.020	0.086	0.090	0.014	0.013	0.02	0.045	0.085	0.20	0.078	0.060	0.044	0.031	0.020	0.021	0.093	0.071
[11,]	0.144	0.122	0.029	0.022	0.055	0.050	0.01	0.011	0.085	0.10	0.155	0.149	0.044	0.031	0.020	0.021	0.031	0.024
[12,]	0.096	0.082	0.086	0.067	0.014	0.013	0.06	0.045	0.021	0.04	0.031	0.149	0.074	0.031	0.020	0.021	0.016	0.012
[13,]	0.024	0.122	0.086	0.045	0.014	0.013	0.01	0.011	0.021	0.02	0.047	0.075	0.148	0.078	0.041	0.021	0.000	0.012
[14,]	0.024	0.020	0.086	0.067	0.055	0.013	0.01	0.011	0.011	0.02	0.016	0.045	0.074	0.156	0.102	0.041	0.031	0.012
[15,]	0.024	0.020	0.014	0.067	0.082	0.050	0.02	0.011	0.011	0.01	0.016	0.015	0.059	0.062	0.245	0.227	0.016	0.012
[16,]	0.024	0.020	0.014	0.011	0.055	0.050	0.08	0.090	0.021	0.02	0.078	0.030	0.177	0.187	0.010	0.247	0.003	0.012
[17,]	0.024	0.020	0.014	0.011	0.014	0.050	0.04	0.022	0.021	0.04	0.062	0.060	0.044	0.047	0.061	0.062	0.467	0.357
[18,]	0.024	0.020	0.014	0.011	0.014	0.013	0.04	0.067	0.085	0.02	0.016	0.015	0.015	0.016	0.020	0.021	0.031	0.357

Joonis 14: α

	päev	kEXn	ksXn	kqXnu	kqXnl	EXn	sXn	qXnu	qXnl
1	0	30	0	30	30	30	0	30	30
8	7	159	12	179	139	140	12	159	120
15	14	1915	172	2197	1633	1549	160	1813	1286
22	21	8651	373	9264	8038	8193	537	9076	7310
29	28	9437	328	9978	8897	12427	393	13073	11780
36	35	2660	189	2971	2349	2376	305	2878	1875
43	42	606	52	692	521	151	28	196	106
50	49	153	19	184	122	8	3	14	3
57	56	42	9	57	27	1	1	2	0
64	63	12	5	20	4	1	1	2	0
71	70	4	3	8	0	1	0	2	0
78	77	2	2	5	-1	NA	NA	NA	NA
85	84	2	1	4	0	NA	NA	NA	NA

Joonis 15: Nakatunude keskmised, standardhälved ja 90% intervallid

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18
1	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.06	0.11	0.17	0.25	0.31	0.33	0.28	0.20	0.12	0.06	0.03
2	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.06	0.11	0.17	0.25	0.31	0.33	0.28	0.20	0.12	0.06	0.03
3	0.00	0.00	0.01	0.01	0.01	0.02	0.04	0.07	0.11	0.18	0.25	0.31	0.33	0.29	0.21	0.12	0.06	0.03
4	0.02	0.02	0.02	0.03	0.03	0.04	0.06	0.09	0.13	0.19	0.27	0.33	0.35	0.30	0.22	0.14	0.08	0.05
5	0.04	0.04	0.04	0.04	0.05	0.06	0.07	0.10	0.15	0.21	0.29	0.35	0.36	0.32	0.24	0.16	0.10	0.07
6	0.05	0.05	0.05	0.05	0.06	0.07	0.08	0.11	0.15	0.22	0.29	0.36	0.37	0.33	0.25	0.17	0.11	0.07
7	0.05	0.05	0.05	0.05	0.06	0.07	0.08	0.11	0.16	0.22	0.30	0.36	0.37	0.33	0.25	0.17	0.11	0.08
8	0.05	0.05	0.05	0.05	0.06	0.06	0.08	0.11	0.15	0.22	0.29	0.36	0.37	0.33	0.25	0.17	0.11	0.07
9	0.04	0.04	0.04	0.05	0.05	0.06	0.08	0.11	0.15	0.21	0.29	0.35	0.37	0.33	0.24	0.16	0.10	0.07
10	0.04	0.04	0.04	0.04	0.05	0.06	0.08	0.10	0.15	0.21	0.29	0.35	0.36	0.32	0.24	0.16	0.10	0.07
11	0.04	0.04	0.04	0.04	0.05	0.05	0.07	0.10	0.14	0.21	0.28	0.35	0.36	0.32	0.24	0.16	0.10	0.06
12	0.03	0.03	0.03	0.04	0.04	0.05	0.07	0.10	0.14	0.20	0.28	0.34	0.36	0.32	0.23	0.15	0.09	0.06
13	0.03	0.03	0.03	0.03	0.04	0.05	0.07	0.09	0.14	0.20	0.28	0.34	0.36	0.31	0.23	0.15	0.09	0.06
14	0.03	0.03	0.03	0.03	0.04	0.05	0.06	0.09	0.14	0.20	0.27	0.34	0.35	0.31	0.23	0.15	0.09	0.05
15	0.03	0.03	0.03	0.03	0.03	0.04	0.06	0.09	0.13	0.20	0.27	0.33	0.35	0.31	0.23	0.15	0.09	0.05
16	0.02	0.02	0.03	0.03	0.03	0.04	0.06	0.09	0.13	0.19	0.27	0.33	0.35	0.31	0.23	0.14	0.08	0.05
17	0.02	0.02	0.02	0.03	0.03	0.04	0.06	0.09	0.13	0.19	0.27	0.33	0.35	0.30	0.22	0.14	0.08	0.05
18	0.02	0.02	0.02	0.02	0.03	0.04	0.06	0.08	0.13	0.19	0.27	0.33	0.35	0.30	0.22	0.14	0.08	0.05

Joonis 16: Nakkusmaatriks

Kasutatud allikad

- Easley, David ja Jon Kleinberg (2010). *Networks, Crowds, and Markets. Reasoning About a Highly Connected World*. New York: Cambridge University Press, lk. 572–576.
- Newman, M.E.J. (2010). *Networks. An Introduction*. New York: Oxford University Press Inc., lk. 628–636.
- Statistikaamet (2022a). *RV021: RAHVASTIK, 1. JAANUAR / Aasta, Vanuserühm ning Sugu. Vanuserühmad 0-85+, mehed ja naised, 2022. aasta seisuga, valitud Saaremaa*.
- (2022b). *RV021: RAHVASTIK, 1. JAANUAR / Aasta, Vanuserühm ning Sugu. Vanuserühmad 0-85+, mehed ja naised, 2022. aasta seisuga, terve Eesti*.
- Sutt, Armand (2022). *Bakalaureuse lõputöö SIR-mudelite koodi git-repository*. URL: <https://github.com/armandsutt/sirmudelloputoo>.

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Armand Sutt,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose Covid andmete põhjal SIR mudeli koostamine, mille juhendaja on Sven Laur PhD, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Armand Sutt

16.08.2022