

KAUR ja TalTech koostatud hundi geneetilise uuringu tulemustele koostatud arvukuse hinnang

Jaanus Remm ja Ants Kaasik, 06.01.2021 / OÜ Rewild

Aastatel 2018–2020 teostati Keskkonnaagentuuri (KAUR) eestvedamisel Mandri-Eesti hundi-asurkonna geneetiline uuring eesmärgiga tuvastada asurkonna suurus ja koosseis¹. Üle Mandri-Eesti koguti hundi elupaikadest jahihooaja vältel, ajavahemikus november 2018 kuni märts 2019, mitteinvasiivsel moel 329 väljaheite proovi. Järgneva laboritöö (proovide ettevalmistamine, DNA eraldamine, järjestamine, analüüs jms) ning andmeanalüüsi teostas Tallinna Tehnikaülikool (TalTech) aastatel 2019–2020.

Vastavalt 17 mikrosatelliit lookuse DNA järjestustele määrati kogutud proovide hulgast hundi väljaheideteks 62 proovi. Ülejäänutest 266-l juhul ei õnnestunud piisava kvaliteediga DNA-d eraldada ning ühel juhul oli tegu koera väljaheitega. Selgus, et sõltuvalt isendite eristamise lähtetingimustest oli 62-st hundi proovist 19 või 12 juhul tegu korduvvaatlustega ning vastavalt 24 või 38 korral ühel korral valimisse sattunud hundiga. Korduvvaatluste arv varieerub johtuvalt sellest, kas sama isendi määranguks lubati DNA järjestuse erinevust mitte üheski (täpne vastavus) või kuni kahes positsioonis. Seega sattus valimisse väljaheiteid kokku 43 või 50-lt hundi isendilt. Võib eeldada, et asurkonnas oli ka isendeid, kes kordagi valimisse ei sattunud.

Eeldades isendite sattumist valimisse juhuslikult, ja seega korduvvaatluste sagedust vastavalt Poissoni jaotusele on võimalik arvutuslikult hinnata valimisse mitte sattunud isendite arvu ning seeläbi asurkonna üldsuurust. Arvutuse aluseks on Poissoni jaotus ja seda iseloomustava parameetri λ seosed üldkogumist (st antud juhul hundi asurkonnast) erineval arvul kordadel (k) valimisse sattunud objektide (antud juhul isendite) osakaaluga (P) vastavalt võrrandile²:

$$P_{X=k} = \lambda^k \times e^{-\lambda} / k! , \quad (1)$$

kus $P_{X=k}$ tähistab objektide valimisse sattumise tõenäosust k hulgal kordadel ning e tähistab Euleri arvu (ca 2,71828). Sellest omakorda järelduvad λ seosed erineval hulgal kordadel valimisse sattunud objektide arvu (N) suhetega:

$$\lambda / (e^\lambda - 1 - \lambda) = N_{k=1} / N_{k \geq 2} \quad (2)$$

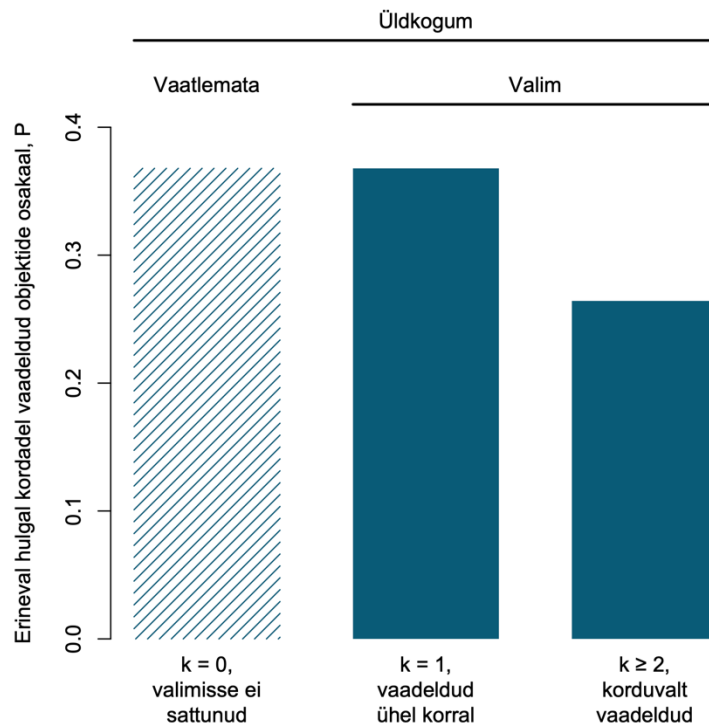
ja

$$\lambda = N_{k=1} / N_{k=0} , \text{ seega } N_{k=0} = N_{k=1} / \lambda . \quad (3)$$

¹ <https://www.keskkonnaagentuur.ee/et/eesmargid-tegevused/projektid/elme/hundi-arvukuse-geneetiline-uuring>

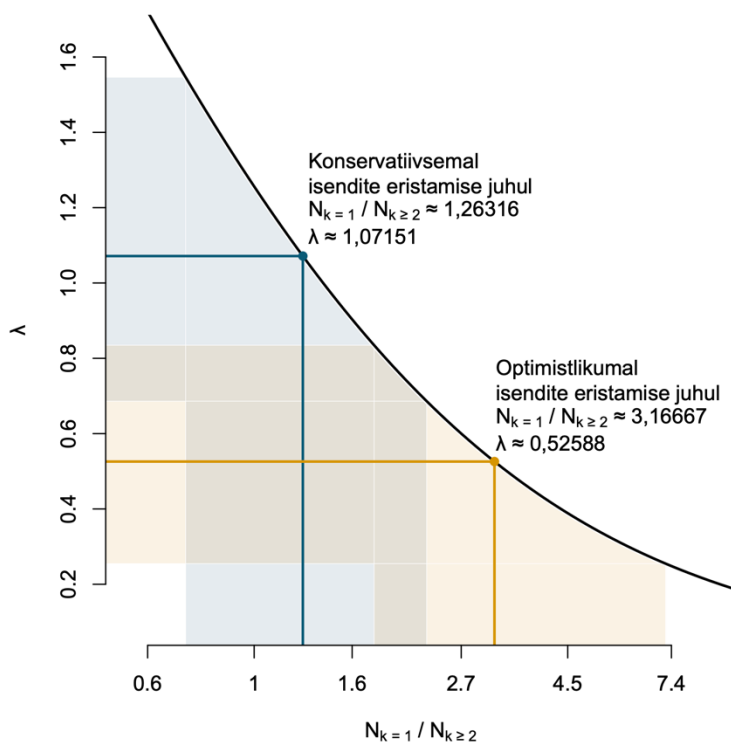
² https://en.wikipedia.org/wiki/Poisson_distribution

Seega saame ühekordsete ja korduvvaatluste arvu, vastavalt $N_{k=1}$ ja $N_{k \geq 2}$ teades tuletada λ väärtuse (võrrand 2) ja sellest omakorda valimisse mitte sattunud isendite hulga ($N_{k=0}$; võrrand 3) ning summeerides leida asurkonna üldarvukuse ($N = N_{k=0} + N_{k=1} + N_{k \geq 2}$; joonis 1).



Joonis 1. Skemaatiline seletus üldkogumi suuruse, st asurkonna üldarvukuse leidmise põhimõttest vastavalt Poissoni jaotusele juhul, kui on teada ühekordsete vaatluste ja korduvvaatluste arv. Vastavalt erineval hulgal kordadel valimisse sattunud objektide (st isendite) jaotusele (täisvärviga tulbad) on võimalik hinnata valimisse mitte sattunud, st vaatlemata objektide osakaal (viirutatud tulp) ning seeläbi kogu üldkogumi (asurkonna) suurus. Näidiseks on kujutatud eri hulgal kordadel vaadeldud objektide osakaalude jaotus juhul, kui Poissoni jaotuse parameeter $\lambda = 1,0$. Pane tähele, et jaotus erineb sõltuvalt λ väärtusest.

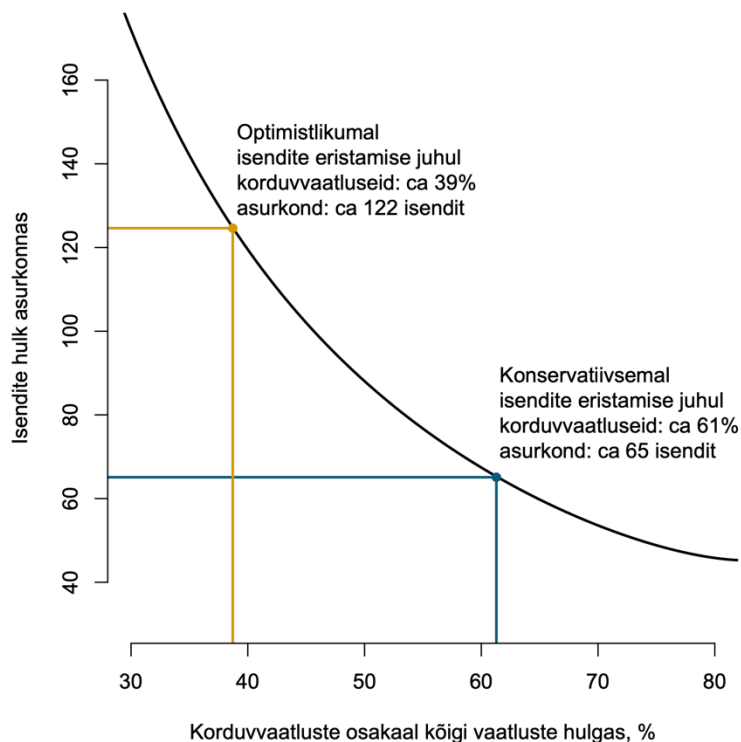
On teada, et kogu valimi suurus oli 62 proovi, mis sisaldas vastavalt isendite eristamise konservatiivsemale hinnangule 19 korduvvaatlust ning vastavalt optimistlikumale hinnangule 12 korduvvaatlust. Seega oli esimesel juhul valimis 24 ühel korral vaadeldud isendit ($N_{k=1}$) ning teisel juhul 38 ühekordset vaatlust. Nii siis saame esimesel juhul ühe- ja mitmekordsete vaatluste suhteks ($N_{k=1} / N_{k \geq 2}$) vastavalt $24 / 19 \approx 1,26$ ja teisel juhul $38 / 12 \approx 3,167$. Edasi järeldub, et konservatiivsemal isendite eristamise juhul on vastav Poissoni jaotuse parameeter $\lambda \approx 1,07$ ning optimistlikumal juhul vastavalt $\lambda \approx 0,53$ (joonis 2). Seega jäi vastavalt võrrandile 3 konservatiivsema isendite eristamise korral valimist välja ($N_{k=0}$) ca 22 isendit ning optimistlikumal juhul ca 72 isendit.



Joonis 2. Graafiline seos Poissoni jaotuse parameetri λ tuletamisest ühekordselt ja korduvalt vaadeldud objektide (isendite) arvude suhtest ($N_{k=1} / N_{k \geq 2}$); must joon vastab võrrandile 2. Varjutatud ala tähistab *Bootstrap* meetodil 100 000 iteratsiooni tulemusena leitud 95% usaldusvahemikku. Rõhthtel on seose visuaalseks lihtsustamiseks logaritmskaalal.

Saadud hinnangu usaldusvahemik leiti juhuslikustatud, nn *Bootstrap* iteratsioonides korrates eelkirjeldatud protseduuri 100 000 korda. Igal iteratsiooni alguses moodustati sama isendite arvu (vastavalt konservatiivsemal juhul 43 ja optimistlikumal juhul 50) juhuvalim sama korduvvaatluste jaotusega ($N_{k=1}$ ja $N_{k \geq 2}$) lähtevalimist, kuid lubades isenditel sattuda iga iteratsiooni valimisse juhuslikult. Saadi 100 000 väärtusest koosnev *Bootstrap* jaotus, mis kirjeldab juhuslikkuse määra korduvvaatluste sageduses ja seega ka saadud $N_{k=0}$ hinnangus. *Bootstrap* jaotusest 2,5% suurimate ja 2,5% väikseimate väärtuste eemaldamisel leiti hundi üldarvukuse hinnangu 95% usaldusvahemik.

Tulemusena selgus, et vastavalt konservatiivsemale hinnangule (st 19 korduvvaatlust) oli 2018/2019 jahihooajal Mandri-Eesti hundi asurkonna suurus 65 isendit (95% usaldusvahemik: 59–78) ning vastavalt optimistlikumale hinnangule (st 12 korduvvaatlust) 122 isendit (95% usaldusvahemik: 96–202; joonis 3).



Joonis 3. Mandri-Eesti hundiasurkonna hinnatud suuruse seos isendite korduvvaatluste sagedusega talvel 2018/2019 kogutud 62 hundi väljaheiteproovide hulgas (must joon). Näidatud on korduvvaatluste sagedus ning sellele vastav arvukuse hinnang juhtudel, kui valimis oli 19 korduvvaatlust (sinine joon) ning 12 korduvvaatlust (kollane joon). Korduvvaatluste hulk erineb sõltuvalt isendite eristamisel lubatud erinevuste hulgast 17 mikrosatelliit lookuse DNA järjestuses.

On oluline panna tähele, et nende hinnangute usaldusvahemik on üsna lai (29–87% võrreldes keskväärtusega) ning arvukus võib olla hiljem muutunud. Tõenäoliselt on eelnevalt leitud arvukuse hinnang mõnevõrra alahinnatud, sest proovid olid kogutud mõnevõrra agregeeritult (võrreldes juhusliku paigutusega) ja tõenäoliselt ei õnnestunud kõiki elupaikasid ühtlaselt katta. Kuna proovid koguti jahihooaja jooksul, siis ei ole võimalik öelda, kas tegu on pigem jahihooaja eelse või järgse arvukuse hinnanguga. Samuti tuleb meeles pidada, et andmeid koguti vaid Mandri-Eestist, kuid mitte Saaremaalt ja Hiiumaalt, mida hundid teadaolevalt vähemalt ajuti siiski asustavad. Lisaks väljaheidetele analüüsiti KAUR ja TalTech uuringus kütitud huntide koeproove ning murtud lammaste karvadelt kogutud süljeproove. Tulemusena eristati eri viisidel kogutud valimi hulgas 88–95 erineva hundi isendi DNA signaal, mis sobib suhteliselt hästi kokku käesoleva arvutuskäiguga saadud arvukuse hinnanguga.