Teoria

- Johdanto simulointiin
- Simuloinnin kulku -- prosessin realisaatioiden tuottaminen
- Satunnaismuuttujan arvonta annetusta jakaumasta
- Tulosten keruu ja analyysi
  - Transienttien tilanteiden simulointi ja tasapainotilanteen simulointi
  - Tilastollinen analyysi ja luottamusvälit
- Varianssinreduktiotekniikoista

Tilastotietojen keruu

- Johdannossa otettiin lähtökohdaksi, että simuloinnin tavoitteena on tarkasteltavan järjestelmän suorituskyvyn arviointi. Simuloimalla siis pyritään arvioimaan jonkin suorituskykyyn liittyvän parametrin arvo α
- Tämä parametri voi liittyä joko järjestelmän **transienttiin** käyttäytymiseen
  - esim. 25 ensimmäisen asiakkaan kokema keskimääräinen odotusaika M/M/1-jonossa tietyllä kuormalla, kun oletetaan, että systeemi on alussa tyhjä

  tai sitten ns. **tasapainotilaan** (steady-state)
  - esim. asiakkaan keskimääräinen odotusaika M/M/1-jonossa tietyllä kuormalla
- Ko. suorituskykyparametri voi toisaalta kuvata tilannetta järjestelmän asiakkaiden kannalta (diskreetisti)
  - esim. Saapuvan asiakkaan keskimäärin näkemä jononpituus M/M/1-jonossa tietyllä kuormalla

  tai sitten systeemin kannalta (jatkuvasti)
  - esim. keskimääräinen jononpituus M/M/1-jonossa tietyllä kuormalla
- Joka tapauksessa yksittäinen simulointiajo tuottaa yhden havainnon, joka jollakin lailla kuvaa arvioitavaa parametria
- Tilastollisten päätelmien tekemiseksi tarvitsemme useita havaintoja (mielellään riippumattomia ja samoin jakautuneita)

Transienttien piirteiden simulointi (1)

- Jos kyseessä on asiakkaiden kokemaan palvelun laatuun liittyvä parametri yksittäinen simulointi päättyy, kun on saatu tietty määrä asiakkaita käsiteltyä
  - esim. oltaessa kiinnostuneita k:n ensimmäisen asiakkaan odotusajasta M/M/1-jonossa, simulointia jatketaan, kunnes viimeinenkin näistä k asiakkasta on saapunut ja päässyt palveluun
- Yksittäisestä simuloinnista saatava havainto X on tässä tapauksessa näiden k:n asiakkaan odotusaikojen W_i keskiarvo ko. simuloinnissa:

$$X = \frac{1}{k} \sum_{i=1}^{k} W_i$$

  - keskeisen raja-arvolauseen perusteella ko. keskiarvoa voidaan pitää ainakin likimain normaalijakaumaa noudattavana (sitä paremmin, mitä enemmän havaintoja)
  - Tilastollisten päätelmien tekemiseksi tarvitsemme riippumattomia ja samoin jakautuneita havaintoja. Näitä saadaan tekemällä useita samanlaisia mutta toisistaan riippumattomia simulointiajoja (toisistaan riippumattomilla satunnaisluvuilla)

Transienttien piirteiden simulointi (2)

- Jos taas kyseessä on systeemin suorituskykyyn liittyvä suure, jota seurataan jatkuvasti, yksittäinen simulointi päättyy ennalta määrätyllä ajanhetkellä T
  - esim. oltaessa kiinnostuneita keskimääräisestä jononpituudesta aikavälillä [0,T], (tapahtumapohjaista) simulointia jatketaan ensimmäiseen hetken T jälkeen tapahtuvaan tapahtumaan asti
- Yksittäisestä simuloinnista saatava havainto X on tässä tapauksessa jononpituuden L(t) aikakeskiarvo yli välin [0,T]

$$X = \frac{1}{T} \int_{0}^{T} L(t)dt$$

  - koska jononpituus ei muutu tapahtumien välillä, ko. integraali on helposti laskettavissa nelikulmioiden summana (huomaa viimeisen tapahtumavälin erityiskäsittely)
  - Tilastollisten päätelmien tekemiseksi tarvitsemme riippumattomia ja samoin jakautuneita havaintoja. Näitä saadaan tekemällä useita samanlaisia mutta toisistaan riippumattomia simulointiajoja (toisistaan riippumattomilla satunnaisluvuilla)

Tasapainotilaan liittyvien piirteiden simulointi (1)

- Tilastotietojen keruu yksittäisestä simuloinnista tapahtuu periaatteessa samalla tavalla kuin transientteja piirteitä simuloitaessa.
- Simuloinnin alussa on kuitenkin ns. lämmittelyvaihe (ennen kuin systeemi on likimain tasapainossa), joka on jätettävä pois kerättävästä datasta.
- Simulointitoistojen tuottamiseksi on tässä tapauksessa ainakin kolme eri tapaa:
  - riippumattomat toistot
  - ns. batch means -menetelmä
  - regeneratiivinen menetelmä

Tasapainotilaan liittyvien piirteiden simulointi (2)

- Riippumattomien toistojen menetelmässä tilastotietojen keruu aloitetaan vasta lämmittelyvaiheen jälkeen.
  - oma ongelmansa on, miten pitkäksi lämmittelyvaihe pitäisi tehdä
- Batch means -menetelmässä tehdään yksi pitkä simulointiajo, joka (keinotekoisesti) jaetaan osiin, joita tietojen keruun kannalta käsitellään omina simulointiajoinaan.
  - tarvitaan vain yksi lämmittelyvaihe, mutta havainnot eivät ole enää täysin riippumattomia
- Regeneratiivisessa menetelmässä vaaditaan, että simuloitava prosessi on regeneroituva. Tällöin kuitenkin saadaan riippumattomia ja samoin jakautuneita havaintoja peräkkäisiltä regeneroitumisjaksoilta.
  - ongelmana on, että jaksojen pituudet voivat satunnaisesti kasvaa hyvinkin pitkiksi
  - esim. G/G/1-jono regeneroituu aina uuden asiakkaan saapuessa tyhjään systeemiin
  - kaikki Markov-prosessit ovat regeneroituvia

Transientin poisto

- Yleensä ollaan kiinnostuneita simuloitavan järjestelmän tasapainotilaan liittyvistä suureista
- Tällöin simuloinnin alkuvaihetta, transienttia, ei tulisi sisällyttää tulosten keruuseen
- Tasapaino on saavutettu silloin, kun systeemin alkutila on ”unohtunut”
  - sillä, mikä alkutila tarkkaanottaen oli, ei ole enää vaikutusta nykyisen tilan jakaumassa
- Transientin poistoon käytetään seuraavia menetelmiä
  - pitkä ajo
  - sopiva initialisointi
  - alkudatan hylkäys
  - batch means
  - regeneratiivinen simulointi

S-38.3148 Tietoverkkojen simulointi / Tulosten keruu ja analyysi

Transientin poisto (jatkoa)

- <u>Pitkä ajo</u>
  - karkea menetelmä
  - jos ajo on kyllin pitkä, alkutransientin vaikutus ”hukkuu” muun datan joukossa
  - vaaditaan hyvin pitkiä ajoja -- hukkaa resursseja
  - vaikea tietää, mikä on riittävän pitkä
- <u>Sopiva initialisointi</u>
  - sen sijaan, että aloitetaan simulointi keinotekoisesta alkutilasta (esim. jonot tyhjiä), käynnistetään systeemi tilasta, joka on lähempänä tasapainoa
    - annetaan eri suureille alkuarvoiksi pitkän aikavälin keskiarvot
    - nämä voidaan likimain tuntea aikaisempien simulointien tai analyyttisten tarkastelujen perusteella
  - tämä <u>vähentää</u> alkutransientin vaikutusta, muttei poista sitä
  - jos tilamuuttujien tasapainojakaumat tunnetaan, alkutransientti voidaan kokonaan poistaa arpomalla muuttujien arvot kyseisistä jakaumista
    - erillinen arvonta jokaisessa toistossa
    - useimmiten jakautumia ei kuitenkaan tunneta

Transientin poisto (jatkoa)

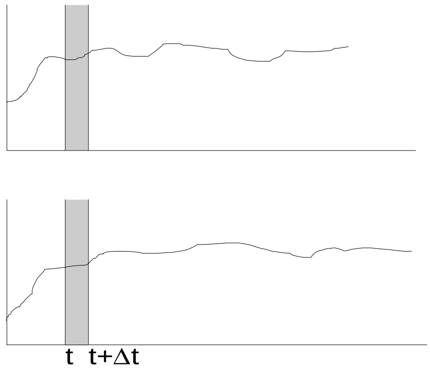
- Alkudatan hylkäys
  - suoraviivainen menetelmä
  - suoritetaan "alkulämmittely" ja kerätään varsinainen data vasta tämän jälkeen
  - ongelma: kuinka pitkä transientti on?
    a) joissakin tapauksissa se tiedetään
    - esimerkiksi tavallisessa häviöjärjestelmässä relaksaatioaika on sama kuin yhteyden (puhelun) keskimääräinen pitoaika
    - karsittava osuus on tällöin n * pitoaika, missä n on luokkaa 3...10
    - alkuarvojen vaikutus on tällöin vähentynyt tekijällä $e^{-10} \approx 10^{-1.3...-4.3}$

    b) yleensä relaksaatioaikaa ei tunneta
    - tällöin voidaan nojautua kokeiluun
    - toistetuissa simuloinneissa kaikista ajoista karsitaan pois samanpituinen osuus
    - annetaan tämän pituuden kasvaa 0:sta ylöspäin ja piirretään mitatun suureen (keski)arvo muulta jaksolta karsitun jakson pituuden funktiona
    - kun keskiarvo ei enää muutu karsitun jakson pituutta kasvatettaessa, on transientti karsittu pois

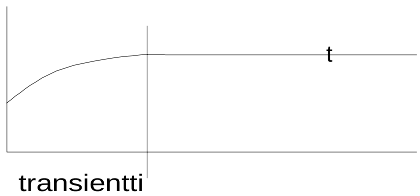
Transientin poisto (jatkoa)

c) Keskiarvo liukuvassa ikkunassa
- tämä on toinen kokeellinen menetelmä transientin keston määräämiseksi
- toistetussa simuloinnissa kustakin ajosta tulokseen kerätään vain tiettyyn (suhteellisen lyhyeen) ikkunaan osuva pätkä realisaatiosta
- lasketaan halutun suureen arvo ikkunan sijainnin funktiona
- keskiarvoistetaan toistojen yli vaihteluiden vähentämiseksi (lyhyen ikkunan sisältä statistiikkaa kertyy vähän ja vaihtelut ovat suuria)
- sijainnin funktiona tarkasteltuna suure yleensä muuttuu alussa ja sitten vakioituu
- kun vakiovaiheeseen on päästy, transientti on ohitettu
- usein transientti on melko lyhyt ja on helppo toimia varman päälle: poistettava jakso voidaan valita vaikkapa kaksinkertaiseksi kokeellisesti määrättyyn transientin kestoon nähden

Transientin keston arviointi liukuvan ikkunan menetelmällä



- Pyritään saamaan käsitys havaittavan suureen “hetkellisen arvon” odotusarvon käyttäytymisestä ajan funktiona
- Odotusarvo määrätään toistokokeiden keskiarvona
- Vaihteluiden vähentämiseksi korvataan hetkellinen arvo kussakin ajossa liukuvan ikkunan sisällä lasketulla keskiarvolla



- Liukuvan ikkunan sijainnin funktiona piirretystä käyrästä voidaan arvioida, milloin transientti loppuu

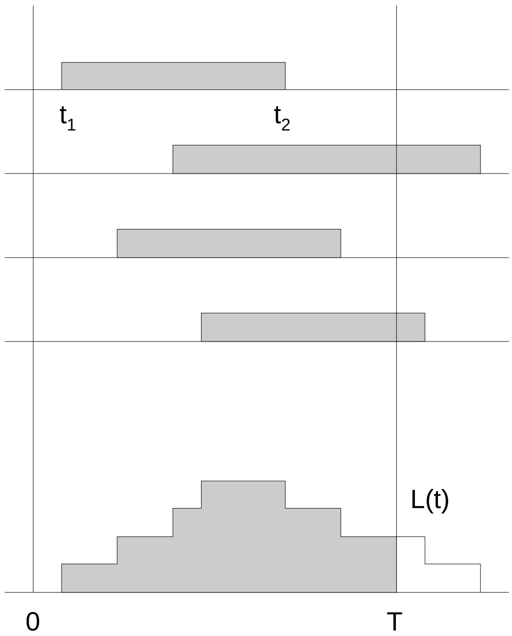
Simuloinnin päättäminen

- Simuloinnin lopetusehdon täyttyessä suoritetaan simuloinnin päättäminen
- Keskeneräisten asioiden käsittelyssä on oltava huolellinen
- Tarkasteltaessa asiakkaan näkökulmaan liittyvää ns. tapahtumapohjaista suuretta
  - tarjottujen kutsujen kokema esto, ylivuotavien pakettien osuus jne

  on otettava huomioon vain ne tapahtumat, jotka on käsitelty loppuun
  - esim. keskimääräinen odotusaika = (niiden asiakkaiden odotusaikojen summa, joiden odotus on loppunut, ts. jotka ovat päässeet palveluun) / (niiden asiakkaiden lukumäärä, joiden odotus on loppunut)
- Tarkasteltaessa systeemin näkökulmaan liittyvää (aikapohjaista) suuretta
  - jononpituus L, aikaosuus jonka systeemi on estotilassa jne

  on keskiarvon laskenta ulotettava simulointijakson T loppuun asti
  - esim. keskimääräinen jononpituus = $\frac{1}{T}\int_{o}^{T} L(t)\,dt$
  - tavallisesti integraalia kerätään vähentämällä saapumishetkellä saapumisaika ja lisäämällä poistumishetkellä poistumisaika
  - erikseen on huomioitava myös ne asiakkaat, jotka ovat sisällä simuloinnin loppuessa; käsitellään ne ikään kuin hetkellä T poistuvina asiakkaina

Jononpituuden aikaintegraalin laskeminen



- Jononpituuden L(t) aikaintegraali muodostuu yksittäisten asiakkaiden jonossa viettämistä ajoista
- Yhden asiakkaan jonossa viettämä aika on lähtöajan t_2 ja tuloajan t_1 erotus t_2-t_1
- L(t):n integraalia voidaan kerätä
  - vähentämällä integraalin arvosta t_1 asiakkaan tullessa jonoon
  - lisäämällä siihen arvo t_2 asiakkaan poistuessa jonosta
- Simuloinnin päättyessä hetkellä T, integraali tulee kerätyksi oikealta ajalta, jos loppuhet-kellä sisällä olevien asiakkaiden katsotaan poistuvan järjestelmästä hetkellä T

S-38.3148 Tietoverkkojen simulointi / Tulosten keruu ja analyysi

Teoria

- Johdanto simulointiin
- Simuloinnin kulku -- prosessin realisaatioiden tuottaminen
- Satunnaismuuttujan arvonta annetusta jakaumasta
- Tulosten keruu ja analyysi
  - Transienttien tilanteiden simulointi ja tasapainotilanteen simulointi
  - Tilastollinen analyysi ja luottamusvälit
- Varianssinreduktiotekniikoista

Parametrien estimointi

- Kuten edellisessä kohdassa todettiin, simuloinnilla pyritään arvioimaan jonkin suorituskykyyn liittyvän parametrin arvo α (esim. asiakkaiden keskimääräinen systeemissäoloaika tai keskimääräinen jononpituus M/M/1-jonossa)
- Yksittäinen simulointi tuottaa kyseisestä parametrista havainnon X_i, joka siis on satunnaismuuttuja. Havaintoa X_i sanotaan **harhattomaksi**, jos $E[X_i] = \alpha$.
- Oletetaan, että olemme saaneet simuloimalla n kpl **riippumattomia** ja **samoin jakautuneita** (i.i.d.) havaintoja. Tällöin niiden **keskiarvo**

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

on parametrin α **harhaton** ja **tarkentuva** estimaattori, sillä

$$E[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \alpha$$

$$D^2[\bar{X}_n] = \frac{1}{n^2}\sum_{i=1}^{n} D^2[X_i] = \frac{1}{n} D^2[X] = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0$$

Esimerkki

- Pyrimme arvioimaan simuloimalla 25:n ensimmäisen asiakkaan keskimääräistä odotusaikaa M/M/1-jonossa kuormalla ρ = 0.9, kun systeemi hetkellä 0 on tyhjä.
- Teoreettinen arvo: α = 2.124
- Kymmenen simulointiajoa ovat tuottaneet seuraavat havainnot X_i (so. keskimääräiset odotusajat kyseisissä simuloinneissa):
  - 1.051, 6.438, 2.646, 0.805, 1.505, 0.546, 2.281, 2.822, 0.414 ja 1.307
- Näiden keskiarvo

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{10}(1.051 + 6.438 + \ldots + 1.307) = 1.982$$

on simulointikokeen antama (25:n ensimmäisen asiakkaan) keskimääräisen odotusajan estimaatti.

Estimaattorin luottamusväli (1)

- Edellä on todettu, että simulointikokeissa saadut havainnot X_i ovat ainakin likimain normaalijakautuneita
- Jos simulointikokeen antamat havainnot X_i noudattaisivat tarkasti normaalijakaumaa $N(\alpha,\sigma^2)$ ja yksittäisen havainnon varianssi $\sigma^2 = D^2[X]$ tunnettaisiin, noudattaisi n:n toiston keskiarvo normaalijakaumaa $N(\alpha,\sigma^2/n)$.
- Tästä saadaan piste-estimaattorina käytetyn havaintojen keskiarvon luottamusväliksi (luottamustasolla 1 - β):

$$\bar{X}_n \pm z_{1-\beta/2} \cdot \frac{\sigma}{\sqrt{n}}$$

missä kerroin z_p tarkoittaa standardin normaalijakauman N(0,1) p-fraktiilia, ts. $P\{Z \le z_p\} = p$, missä $Z \sim N(0,1)$

- Tulkinta: estimoitava parametri α on tn:llä 1 - β kyseisellä välillä.
- Esimerkiksi 95%:n luottamustasoa vastaa kerroin $z_{0.975} \approx 1.960$

Estimaattorin luottamusväli (2)

- Yleensä emme kuitenkaan tunne yksittäisen havainnon varianssia $\sigma^2 = D^2[X]$.
- Sitä voidaan kuitenkin puolestaan estimoida ns. otosvarianssilla

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

joka on (riippumattomien ja samoin jakautuneiden havaintojen tapauksessa) varianssin harhaton estimaattori.

- Otoshajonta on otosvarianssin neliöjuuri:

$$S_n = \sqrt{S_n^2}$$

Estimaattorin luottamusväli (3)

- Jos simulointikokeen antamat havainnot X_i noudattaisivat tarkasti normaalijakaumaa $N(\alpha,\sigma^2)$, noudattaisi otoshajonnalla sopivasti normeerattu otoskeskiarvo ns. Studentin t-jakaumaa vapausastein n-1.
- Tästä saadaan piste-estimaattorina käytetyn havaintojen keskiarvon luottamusväliksi (luottamustasolla $1 - \beta$):

$$\overline{X}_n \pm t_{n-1,1-\beta/2} \cdot \frac{S_n}{\sqrt{n}}$$

  missä kerroin $t_{n-1,p}$ tarkoittaa t-jakauman (vapausastein n-1) p-fraktiilia, ts. $P\{T \leq t_{n-1,p}\} = p$, missä T noudattaa ko. t-jakaumaa
- Tulkinta: estimoitava parametri α on tn:llä $1 - \beta$ kyseisellä välillä.
- Esimerkiksi 95%:n luottamustasoa vastaa
  - 10 havainnon tapauksessa kerroin $t_{9,0.975} \approx 2.262$ ja
  - 101 havainnon tapauksessa kerroin $t_{100,0.975} \approx 1.984$

Esimerkki (jatkoa)

- Pyrimme arvioimaan simuloimalla 25:n ensimmäisen asiakkaan keskimääräistä odotusaikaa M/M/1-jonossa kuormalla $\rho = 0.9$, kun systeemi hetkellä 0 on tyhjä.
- Teoreettinen arvo: $\alpha = 2.124$
- Kymmenen simulointiajoa ovat tuottaneet seuraavat havainnot X_i (so. keskimääräiset systeemissäoloajat kyseisissä simuloinneissa):
  - 1.051, 6.438, 2.646, 0.805, 1.505, 0.546, 2.281, 2.822, 0.414 ja 1.307
- Otoskeskiarvoksi saatiin 1.982 ja otoshajonnaksi tulee

$$S_n = \sqrt{\tfrac{1}{9}((1.051 - 1.982)^2 + \ldots + (1.307 - 1.982)^2)} = 1.781$$

- Simulointikokeen antama 25:n ensimmäisen asiakkaan keskimääräisen odotusajan piste-estimaatin luottamusväli 95%:n luottamustasolla on siis

$$\overline{X}_n \pm t_{n-1,1-\beta/2} \cdot \frac{S_n}{\sqrt{n}} = 1.982 \pm 2.262 \cdot \frac{1.781}{\sqrt{10}} = 1.982 \pm 1.274$$

Havaintoja

- Simulointikokeen tulos tarkentuu (so. piste-estimaatin luottamusväli kapenee), kun
  - simulointitoistojen eli riippumattomien havaintojen lukumäärää n kasvatetaan
  - yksittäisen havainnon varianssia pienennetään (esim. ajamalla pitempiä yksittäisiä simulointiajoja tai muilla ns. varianssin reduktiomenetelmillä)
  - Jos on annettu haluttu simulointitulosten suhteellinen tarkkuus (so. luottamusvälin puolikkaan suhde otoskeskiarvoon), voidaan dynaamisesti seurata, kuinka monta riippumatonta simulointitoistoa on tehtävä ko. tavoitteeseen pääsemiseksi