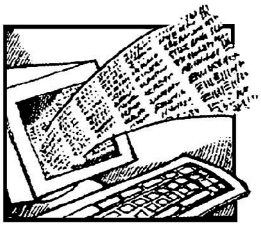


Odottelua pysäkillä

Juha Karvanen ja Antti Luoto
Matematiikan ja tilastotieteen laitos, Jyväskylän yliopisto

Kiirehdit bussipysäkille. Saapuessasi paikalle kello on 8.04. Arvelet, että bussi ei ole vielä mennyt, mutta et ole täysin varma asiasta. Bussi lähtee päätepysäkiltään klo 8.00 ja tälle pysäkille matka kestää yleensä noin kahdeksan minuuttia. Jäät odottelemaan bussia, jonka uskot saapuvan pian. Kymmenen minuutin odottelun jälkeen mieleesi nousee aavistus siitä, että bussi on sittenkin ehtinyt jo ohittaa pysäkin ennen kuin tulit paikalle. Odoteltuasi vielä viisi minuuttia olet hyvin varma siitä, että olet myöhästynyt bussista. Seuraava bussi lähtee klo 8.30 ja siihen ainakin ehdit hyvin. Odottelu pysäkillä jatkuu.

Edellä esitettyä arkipäivän tilannetta voi tarkastella todennäköisyyslaskennan keinoin. Tiedämme siis, että bussi lähtee klo 8.00, matkustaja saapuu pysäkille klo 8.04 ja seuraava bussi lähtee klo 8.30. Kiinnostuksen kohteena oleva kysymys kuuluu: ”Mikä on jäljellä oleva keskimääräinen odotteluaika, jos matkustaja on odottanut pysäkillä jo s minuuttia?” Täsmällisempi todennäköisyyslaskennan termi keskimääräiselle odotteluajalle on odotteluajan odotusarvo.

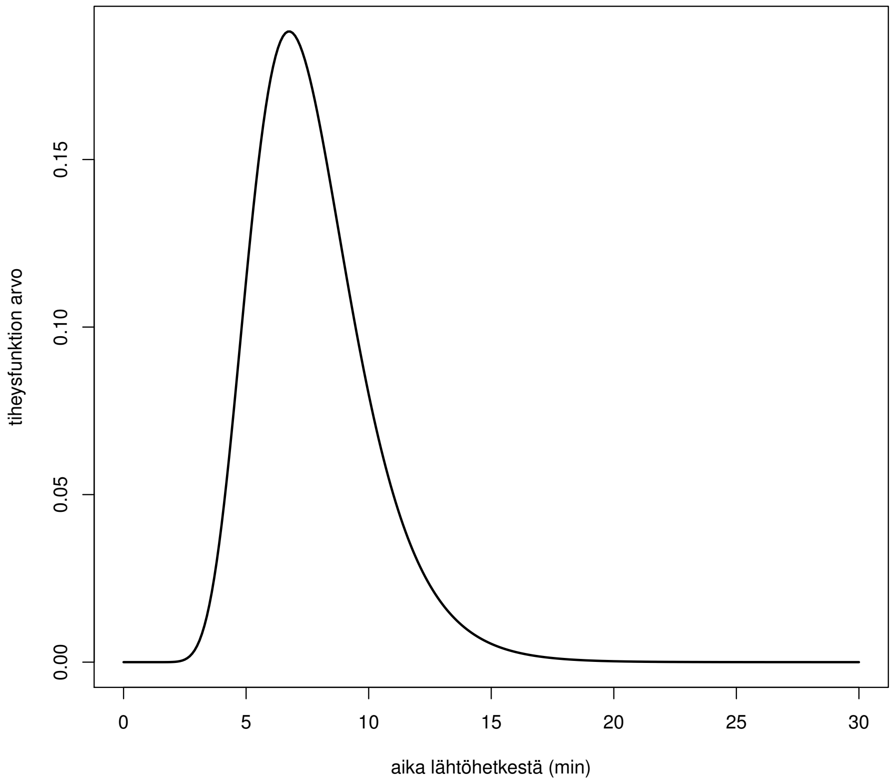
Jotta tehtävän voisi ratkaista, tarvitsemme tiedon ajasta T, joka bussilla kuluu päätepysäkiltä odottelupysäkille. Tämä aika ei ole vakio, vaan se vaihtelee riippuen mm. matkustajien ja muun liikenteen määrästä. Keräämällä systemaattisesti havaintoja bussin saapumisajoista saamme tietoa satunnaismuuttujan T jakaumasta. Havaintoaineiston perusteella matka-ajan jakaumalle voi sovittaa jonkin sopivan mallin, kuten esimerkiksi gamma-jakauman tai log-normaalijakauman.

Oletamme nyt, että tiedämme tämän matka-ajan jakauman olevan log-normaalijakauma. Sanotaan, että T on log-normaalijakautunut parametrein μ ja $\sigma > 0$, $T \sim \log \mathcal{N}(\mu, \sigma)$, mikäli tiheysfunktio on

$$f_T(t) = \begin{cases} \frac{1}{t\sqrt{2\pi\sigma^2}} e^{-\frac{(\log t - \mu)^2}{2\sigma^2}}, & t > 0 \\ 0, & \text{muulloin} \end{cases}$$

ja kertymäfunktio on

$$F_T(t) = \begin{cases} \Phi\left(\frac{\log t - \mu}{\sigma}\right) = \\ \qquad \int_{-\infty}^{\frac{\log t - \mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy, & t > 0 \\ 0, & \text{muulloin,} \end{cases} \tag{1}$$



Kuva 1: Bussin saapumisaikaa kuvaavan log-normaalijakauman tiheysfunktio.

missä Φ on standardinormaalijakauman $\mathcal{N}(0,1)$ kertymäfunktio. Tyypillinen tapa johtaa log-normaalijakauma on tarkastella muunnosta $T = g(X) = e^X$, missä satunnaismuuttuja X noudattaa normaalijakaumaa, $X \sim \mathcal{N}(\mu, \sigma)$, jolloin tiheysfunktio f_T ratkeaa muuttujanvaihdon avulla. Siis $T \sim \log\mathcal{N}(\mu, \sigma)$ mikäli $\log T \sim \mathcal{N}(\mu, \sigma)$. Parametri μ kuvaa jakauman sijaintia ja $\sigma > 0$ on skaalaparametri. Jatkuvan, vain positiivisia arvoja saavan satunnaismuuttujan odotusarvo ja varianssi määritellään kaavoilla

$$\mathbb{E}[T] = \int_0^\infty t f_T(t) dt$$

ja

$$\mathbf{Var}[T] = \int_0^\infty (t - \mathbb{E}[T])^2 f_T(t) dt.$$

Log-normaalijakauman $\log\mathcal{N}(\mu, \sigma)$ tapauksessa odotusarvoksi ja varianssiksi saadaan

$$\mathbb{E}[T] = e^{\mu + \frac{1}{2}\sigma^2} \quad \text{ja} \quad \mathbf{Var}[T] = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}.$$

Kun parametrien arvoiksi kiinnitetään $\mu = 2$ ja $\sigma = 0.3$, saadaan bussin matka-ajan odotusarvoksi $e^{2 + \frac{1}{2}\cdot 0.3^2} \approx 7.73$ minuuttia, joka pyöristettynä on edellä mainittu kahdeksan minuuttia. Jakauman tiheysfunktion kuvaaja on esitetty kuvassa 1. Kuvasta nähdään, että matka-ajan jakauma on vino. Jakauman moodi eli tiheysfunktion huippukohta on $e^{2 - 0.3^2} \approx 6.75$ minuuttia. Kolmas jakauman sijaintia kuvaava tunnusluku on mediaani, jonka arvo $e^2 \approx 7.39$ minuuttia. Mediaanin kohdalla kertymäfunktio saa arvon 0.5. Yksihuippuisen symmetrisen jakauman tapauksessa näiden kolmen tunnusluvun arvot olisivat samat, mutta koska log-normaalijakauma on vino jakauma, kaikki kolme tunnuslukua eroavat toisistaan.

Bussin matka-ajan jakaumamallin ollessa selvillä voimme laskea sen perusteella meitä kiinnostavia todennäköisyyksiä. Numeerisiin laskutoimituksiin voi käyttää esimerkiksi avoimen lähdekoodin R-ohjelmistoa, joka on saatavilla osoitteesta www.r-project.org. Todennäköisyys sille, että bussi ohittaa pysäkin ennen kuin kello on 8.04 vastaa todennäköisyyttä, että matka-aika on alle neljä minuuttia. Tämä lasketaan R-ohjelman komennolla `plnorm(4,meanlog=2,sdlog=0.3)` ja tulokseksi saadaan $\mathbb{P}(T < 4) \approx 0.02$. Bussista myöhästyminen on siis mahdollista, mutta ei kovin todennäköistä, kun pysäkille saavutaan klo 8.04.

Vastaavalla tavalla voidaan laskea todennäköisyys sille, että matka-aika on enemmän kuin $4 + s$ minuuttia. Tämän voi tehdä R-komennolla `1-plnorm(4+s,meanlog=2,sdlog=0.3)`. Koska tehtävän tilanteessa tiedämme, että bussi ei ole saapunut sinä aikana, kun matkustaja odottelee pysäkillä, täytyy päteä joko $T < 4$ tai $T > 4 + s$. Näistä jälkimmäisen mahdollisuuden todennäköisyys pienenee, kun odotteluaika s kasvaa.

Jos klo 8.00 lähtenyt bussi on jo mennyt, pysäkillä odottelua kestää siihen saakka, kunnes klo 8.30 lähtevä bussi saapuu. Odotteluajan $W^{(s)}$ odotusarvoksi saadaan tällöin $\mathbb{E}(W^{(s)}) = 30 - (4+s) + \mathbb{E}(T) \approx 33.8 - s$. Tarkastelussa rajoitutaan tapauksiin, joissa pysäkillä on odoteltu korkeintaan klo 8.30 asti eli $s \leq 26$.

Jos klo 8.00 lähtenyt bussi ei ole vielä mennyt, jäljellä olevan odotteluajan odotusarvo on ehdollinen odotusarvo $\mathbb{E}[T \mid T > s+4]$. Tämä odotusarvo voidaan laskea analyyttisesti tai simuloimalla. Analyyttinen ratkaisu edellyttää integrointia, sopivan muuttujanvaihdon ja neliöksi täydentämistä. Lopuksi päädytään tulokseen, joka on ilmaistavissa normaalijakauman kertymäfunktion avulla. Simulointiratkaisu taas edellyttää, että klo 8.00 ja klo 8.30 lähtevien bussien saapumisaikoja generoidaan esimerkiksi R-ohjelmiston avulla suuri joukko, jonka avulla odotteluajan odotusarvo voidaan määrittää likimääräisesti. Tulos on sitä tarkempi, mitä enemmän saapumisaikoja on generoitu.

Ongelman ratkaiseminen analyyttisesti

Ongelma keskimääräisen odotteluajan ratkaisemiseksi voidaan muotoilla matemaattisesti todennäköisyyslaskennan käsitteiden avulla. Täsmällisyyden vuoksi esityksessä käytetään todennäköisyysavaruuden ja Borel-mitallisuuden käsitteitä, joiden syvällinen ymmärtäminen ei kuitenkaan ole välttämätöntä esityksen seuraamiseksi. Aiheesta kiinnostunut lukija voi tutustua todennäköisyyslaskennan teoreettisiin perusteisiin tarkemmin esimerkiksi Wikipedia-sivulla ”Todennäköisyysteoria”.

Olkoot T_1, T_2 riippumattomia $\log\mathcal{N}(\mu,\sigma)$-jakautuneita satunnaismuuttujia abstraktilla todennäköisyysavaruudella $(\Omega, \mathcal{F}, \mathbb{P})$. Satunnaismuuttujat T_1 ja T_2 tulkitaan klo 8.00 (bussi 1) ja klo 8.30 (bussi 2) lähtevien bussien matka-ajoiksi, ja nämä oletetaan siis toisistaan riippumattomiksi. Tiedämme, että matka-ajalle T_1 pätee joko $T_1 < 4$ tai $T_1 > 4+s$. Mikäli $T_1 > 4+s$, seuraavaksi pysäkille saapuu bussi 1. Tässä tapauksessa joudutaan odottelemaan vielä $T_1 - (4+s)$ minuuttia. Mikäli puolestaan $T_1 < 4$, pysäkille saapuu seuraavaksi bussi numero 2, joka saapuu paikalle $T_2 + 30$ minuuttia yli kahdeksan. Odoteltavaa jää siten $T_2 + 30 - (4+s) = T_2 + 26 - s$ minuuttia. On täten luontevaa määritellä jäljellä oleva odotusaika $W^{(s)}$ deterministisen ajan s ja satunnaisten aikojen T_1 ja T_2 funktiona seuraavalla tavalla: annetulle $s \in [0, 26]$ asetetaan

$$\begin{aligned} W^{(s)}(\omega) &:= f(s, T_1(\omega), T_2(\omega)) \\ &:= \begin{cases} T_2(\omega) + 30 - (4+s), & \text{jos } T_1(\omega) < 4 \\ 0, & \text{jos } 4 \leq T_1(\omega) \leq 4+s \\ T_1(\omega) - (4+s), & \text{jos } T_1(\omega) > 4+s. \end{cases} \end{aligned}$$

Näin määriteltynä $W^{(s)}$ on satunnaismuuttuja avaruudella $(\Omega, \mathcal{F}, \mathbb{P})$, mikä seuraa kuvauksen $(t_1, t_2) \mapsto f(s, t_1, t_2)$ Borel-mitallisuudesta. Koska ongelman kannalta on merkityksetöntä, mitä arvoja $W^{(s)}$ saa joukossa $\{4 \leq T_1 \leq 4+s\}$, valittiin yksinkertaisuuden vuoksi vakioarvo 0.

Näillä merkinnöin keskimääräinen odotteluaika, josta käytetään merkintää $e(s)$, on tarkalleen ottaen muuttujan $W^{(s)}$ ehdollinen odotusarvo ehdolla tapahtuma $\{T_1 \notin [4, 4+s]\}$. Koska $\mathbb{P}(T_1 \notin [4, 4+s]) > 0$, edellä mainittu ehdollinen odotusarvo määritellään klassista ehdollista todennäköisyyttä muistuttavalla kaavalla

$$e(s) = \mathbb{E}\big[W^{(s)} \big| T_1 \notin [4, 4+s]\big] := \frac{\mathbb{E}\big[W^{(s)} \mathbf{1}_{\{T_1 \notin [4,4+s]\}}\big]}{\mathbb{P}(T_1 \notin [4, 4+s])},$$

missä $\mathbf{1}_A$ tarkoittaa indikaattorifunktiota,

$$\mathbf{1}_{\{t_1 \in A\}} = \begin{cases} 1, & \text{jos } t_1 \text{ kuuluu joukkoon A} \\ 0 & \text{muuten.} \end{cases}$$

Todennäköisyyslaskennan laskusääntöjä noudattaen voidaan kirjoittaa

$$\begin{aligned} \mathbb{P}(T_1 \notin [4, 4+s]) &= \mathbb{P}(T_1 \leq 4) + \mathbb{P}(T_1 > 4+s) \\ &= F_T(4) + 1 - F_T(4+s), \end{aligned}$$

missä F_T on kaavassa (1) määritelty log-normaalijakauman kertymäfunktio. Edelleen, odotusarvon lineaarisuuden ja muuttujien T_1 ja T_2 riippumattomuuden nojalla pätee toinen yhtäsuuruus

$$\begin{aligned} &\mathbb{E}\big[W^{(s)} \mathbf{1}_{\{T_1 \notin [4,4+s]\}}\big] \\ &= \mathbb{E}\big[W^{(s)} \mathbf{1}_{\{T_1 < 4\}}\big] + \mathbb{E}\big[W^{(s)} \mathbf{1}_{\{T_1 > 4+s\}}\big] \\ &= \mathbb{E}\big[\big(T_2 + 30 - (4+s)\big) \mathbf{1}_{\{T_1 < 4\}}\big] \\ &\qquad + \mathbb{E}\big[\big(T_1 - (4+s)\big) \mathbf{1}_{\{T_1 > 4+s\}}\big] \\ &= \mathbb{E}\big[T_2 + 30 - (4+s)\big] \mathbb{P}(T_1 < 4) \\ &\qquad + \mathbb{E}\big[T_1 \mathbf{1}_{\{T_1 > 4+s\}}\big] - (4+s)\mathbb{P}(T_1 > 4+s) \\ &= \big(\mathbb{E}[T_2] + 26 - s\big) F_T(4) \\ &\qquad + \mathbb{E}\big[T_1 \mathbf{1}_{\{T_1 > 4+s\}}\big] - (4+s)\big(1 - F_T(4+s)\big). \end{aligned}$$

Näin ollen

$$e(s) = \frac{M}{F_T(4) + 1 - F_T(4+s)}, \tag{2}$$

missä

$$\begin{aligned} M = \big(\mathbb{E}[T_2] + 26 - s\big) F_T(4) + \mathbb{E}\big[T_1 \mathbf{1}_{\{T_1 > 4+s\}}\big] \\ - (4+s)\big(1 - F(4+s)\big). \end{aligned}$$

Tekemällä muuttujanvaihto $t = e^y$ nähdään, että

$$\begin{aligned} &\mathbb{E}\big[T_1 \mathbf{1}_{\{T_1 > 4+s\}}\big] \\ &\qquad = \int_{4+s}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log t - \mu)^2}{2\sigma^2}} dt \\ &\qquad = \int_{\log(4+s)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y-\mu)^2 + 2\sigma^2 y}{2\sigma^2}} dy. \end{aligned} \tag{3}$$

Neliööntäydennys muuttujan y suhteen puolestaan antaa

$$-(y-\mu)^2 + 2\sigma^2 y = -\big(y - (\mu + \sigma^2)\big)^2 + 2\mu\sigma^2 + \sigma^4,$$

joten kaavan (3) oikea puoli yksinkertaistuu lopulta muuttujanvaihtoa $x = \frac{y}{\sigma} - \frac{(\mu+\sigma^2)}{\sigma}$ ja symmetriaa $\Phi(x) = \Phi(-x)$ soveltaen muotoon

$$\begin{aligned}
&\int_{\log(4+s)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y-\mu)^2+2\sigma^2 y}{2\sigma^2}} dy \\
&= e^{\mu+\sigma^2/2} \int_{\log(4+s)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y-(\mu+\sigma))^2}{2\sigma^2}} dy \\
&= e^{\mu+\sigma^2/2} \int_{\log(4+s)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y/\sigma-(\mu+\sigma^2)/\sigma)^2}{2}} dy \\
&= e^{\mu+\sigma^2/2} \int_{\frac{\log(4+s)-(\mu+\sigma^2)}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= e^{\mu+\sigma^2/2} \Phi\Big(\frac{-\log(4+s)+\mu+\sigma^2}{\sigma}\Big).
\end{aligned}$$

Siis $\mathbb{E}[T_1 \mathbf{1}_{\{T_1>4+s\}}] = e^{\mu+\sigma^2/2}\Phi\Big(\frac{-\log(4+s)+\mu+\sigma^2}{\sigma}\Big)$. Koska lisäksi $\mathbb{E}[T_2] = e^{\mu+\frac{1}{2}\sigma^2}$, kaavaan (2) sijoittamalla saadaan lopulta

$$e(s) = \frac{N}{F_T(4) + 1 - F_T(4+s)},$$

missä

$$\begin{aligned}
N = \Big[e^{\mu+\sigma^2/2} + 26 - s\Big] F_T(4) \\
+ e^{\mu+\sigma^2/2}\Phi\Big(\frac{-\log(4+s)+\mu+\sigma^2}{\sigma}\Big) \\
- (4+s)\big(1 - F_T(4+s)\big).
\end{aligned}$$

Annetuilla parametrien arvoilla $\mu = 2$ ja $\sigma = 0.3$ keskimääräiseksi odotusajaksi saadaan siten esimerkiksi

$$e(s) = \begin{cases} 3.24, & s = 3 \\ 6.27, & s = 7 \\ 18.07, & s = 15 \\ 8.73, & s = 25. \end{cases}$$

Ongelman ratkaiseminen simuloimalla

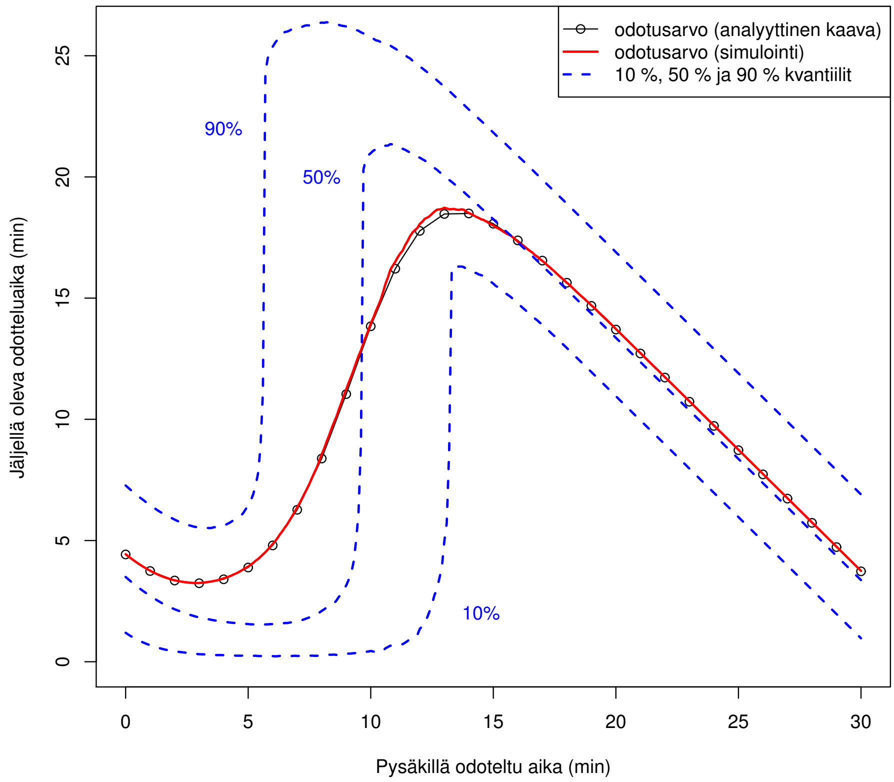
Samaan tulokseen voidaan päästä myös simuloimalla. Tällöin log-normaalijakaumasta generoidaan suuri määrä saapumisaikoja klo 8.00 ja 8.30 lähteville busseille. Olkoon T_1 ensimmäisen ja T_2 toisen bussin matka-aika pysäkille. Tarkastelusta poistetaan ne tapaukset, joissa ensimmäinen bussi saapuu aikavälillä $[4, 4+s]$. Jokaiselle jäljelle jääneelle n tapaukselle lasketaan odotteluaika generoitujen saapumisaikojen perusteella. Jos $T_1 > 4+s$ eli ensimmäinen bussi ei ole vielä mennyt, jäljellä oleva odotteluaika on $W^{(s)} = T_1 - 4 - s$. Jos taas $T_1 < 4$ eli ensimmäinen bussi on jo mennyt, jäljellä oleva odotteluaika on $W^{(s)} = 30 + T_2 - 4 - s$. Simuloiduista tapauksista laskettu muuttujan $W^{(s)}$ keskiarvo estimoi odotteluajan odotusarvoa. Tuloksen numeerinen tarkkuus riippuu tapausten lukumäärästä n, joka puolestaan riippuu simuloinnissa käytettyjen toistojen lukumäärästä.

Simuloinnin voi toteuttaa esimerkiksi alla olevalla R-koodilla:

```
n <- 100000 #toistojen lkm
mu <- 2
sigma <- 0.3
r <- 4
interval <- 30
T1 <- rlnorm(n,mu,sigma)
T2 <- interval + rlnorm(n,mu,sigma)
svec <- 0:30 #odotteluaikojen vektori
wn <- rep(NA,length(svec)) #tarkasteltavien
                            #tapausten lkm
wmean <- rep(NA,length(svec)) #odotteluajan
                               #odotusarvo
wmedian <- rep(NA,length(svec)) #odotteluajan
                                 #mediaani
wq10 <- rep(NA,length(svec)) #odotteluajan
                              #10 % kvantiili
wq90 <- rep(NA,length(svec)) #odotteluajan
                              #90 % kvantiili
for(i in 1:length(svec)) #tehdään laskenta
             #erikseen jokaiselle s:n arvolle
{
 s <- svec[i]
 valid <- (T1>(s+r) | T1<r) #indikaattori:
                #kuuluuko tarkastelujoukkoon
 wn[i] <- sum(valid)
 intime <- (r < T1[valid]) #indikaattori:
                  #bussi ei ole vielä mennyt
 wtime <- intime*(T1[valid]-r-s)+(!intime)*
                              (T2[valid]-r-s)
 wmean[i] <- mean(wtime)
 wmedian[i] <- median(wtime)
 wq10[i] <- quantile(wtime,0.1)
 wq90[i] <- quantile(wtime,0.9)
}

#Vertaillaan analyyttisesti johdettuun
#odotusarvoon, jonka arvoja voi laskea
#seuraavalla funktiolla
waitingtime <- function(s,alpha=2,sigma=0.3,
                        r=4,interval=30)
{
 sr <- 1-plnorm(s+r,alpha,sigma)
 fr <- plnorm(r,alpha,sigma)
 E1 <- exp(alpha+sigma^2/2)/sr*
      pnorm((-log(s+r)+alpha+sigma^2)/sigma)-
      (s+r)
 E2 <- interval-r-s+exp(alpha+sigma^2/2)
 E <- sr/(sr+fr)*E1+fr/(sr+fr)*E2
 return(E)
}

#Piirretään tuloksista kuva
```



Kuva 2: Jäljellä olevan odotteluajan odotusarvo, mediaani sekä 10 % ja 90 % kvantiilit pysäkillä odotellun ajan funktiona.

```
s <- 0:30
plot(s,waitingtime(s,r=4),type=''o'',
  ylim=c(0,26),
  xlab=''Pysäkillä odoteltu aika (min)'',
  ylab=''Jäljellä oleva odotteluaika (min)'')
lines(svec,wmean,col=''red'',lwd=2)
lines(svec,wmedian,col=''blue'',lty=2,lwd=2)
lines(svec,wq10,col=''blue'',lty=2,lwd=2)
lines(svec,wq90,col=''blue'',lty=2,lwd=2)
legend(''topright'',
 legend=c(''odotusarvo (analyyttinen kaava)'',
          ''odotusarvo (simulointi)'',
          ''10 %, 50 % ja 90 % kvantiilit''),
 lty=c(1,1,2),lwd=c(1,2,2),
 pch=c(1,NA,NA),
 col=c(''black'',''red'',''blue''))
text(4,22,''90%'',col=''blue'')
text(8,20,''50%'',col=''blue'')
text(14.5,2,''10%'',col=''blue'')
```

Odotteluajan odotusarvon lisäksi R-koodissa estimoidaan myös odotteluajan mediaani, 10 % ja 90 % kvantiilit. Mediaani eli 50 % kvantiili tarkoittaa havaintoa, jota pienempiä on 50 % havainnoista. Vastaavasti 10 % ja 90 % kvantiilit tarkoittavat havaintoja, joita pienempiä on 10 % ja 90 % havainnoista. Kvantiilien ratkaiseminen ei onnistu helposti analyyttisesti, mutta simuloinnissa ne saadaan laskettua yhden koodirivin lisäyksellä.

Tulokset on esitetty graafisesti kuvassa 2, joka on luotu edellä esitetyllä R-koodilla. Kuvasta nähdään, että analyyttisesti laskettu odotusarvo ja simuloimalla laskettu odotusarvo ovat hyvin lähellä toisiaan. Aluksi odotteluajan odotusarvo pienenee, kun s kasvaa, mutta kolmen minuutin odottelun jälkeen (klo 8.07) odotusarvo alkaa kasvaa. Maksimi saavutetaan 14 minuutin odottelun jälkeen (klo 8.18) , jonka jälkeen odotusarvo alkaa pienentyä lähes lineaarisena funktiona. Mediaanin ja muiden kvantiilien hyppäykset ovat seurausta jakauman kaksihuippuisuudesta. Todennäköisyys, että ensimmäinen bussi on jo mennyt ylittää 10 % noin kuuden minuutin odottelun jälkeen (klo 8.10), 50 % noin kymmenen minuutin odottelun jälkeen (klo 8.14) klo ja 90 % noin 13 minuutin odottelun jälkeen (klo 8.17).

Esimerkin tarkoituksena on ollut havainnollistaa sitä, kuinka läheisesti todennäköisyyslaskenta liittyy arkipäivän tapahtumiin. Lisäksi tavoitteena on ollut esittää, kuinka jakaumiin liittyviä tehtäviä voidaan ratkaista sekä analyyttisesti että simuloimalla. Analyyttinen ratkaiseminen ja simulointi täydentävät toinen toistaan. Ratkaisu, joka saadaan kahdella toisistaan riippumattomalla tavalla, on todennäköisemmin oikein kuin vain yhdellä tavalla saatu ratkaisu.

(Esimerkkiä on käytetty harjoitustehtävänä Jyväskylän yliopiston kurssilla Elinaikamallit.)