Madelon de Kemp
Korteweg-de Vries Instituut voor Wiskunde
Universiteit van Amsterdam
m.a.dekemp@uva.nl

Onderzoek

Inplannen van afspraken voor een groot aantal patiënten

Voor haar masterscriptie heeft Madelon de Kemp onderzoek gedaan naar het optimaal inplannen van afspraken. Hiervoor heeft zij vorig jaar november de ASML afstudeerprijs ontvangen, uitgereikt door de Koninklijke Hollandsche Maatschappij der Wetenschappen (KHMW). In dit artikel vertelt zij iets over haar onderzoek.

Stel we willen graag afspraken inplannen bij bijvoorbeeld een dokter. Aan de ene kant willen we dat patiënten niet te lang hoeven wachten. Aan de andere kant willen we dat de dokter niet teveel tijd onbenut laat tussen de behandeling van twee patiënten in. Hiervoor moeten we twee zaken van tevoren bepalen: hoeveel tijd er voor iedere afspraak wordt ingepland, en in welke volgorde de dokter de verschillende patiënten ziet.

De tijd die een behandeling duurt is niet van tevoren bekend. In plaats daarvan zullen we aannemen dat de behandeltijd een stochast is. We nemen aan dat de behandeltijden van verschillende patiënten onafhankelijk zijn. We zullen ook aannemen dat we voor iedere patiënt de verwachting en variantie van de behandeltijd kennen, bijvoorbeeld door vorige metingen, maar niet de volledige verdeling. Onze planning moet dan goed presteren voor alle mogelijke verdelingen met deze verwachting en variantie.

Doelfunctie

Voordat we op zoek kunnen gaan naar een optimale planning, zullen we eerst precies moeten maken wat we bedoelen met optimaal. Noteer de wachttijd van patiënt i met W_i. We willen graag dat $\mathbb{E}W_i$, de verwachte wachttijd van patiënt i, klein is. Noteer met I_i de tijd die de dokter onbenut laat tussen patiënt $i-1$ en patiënt i. We willen dan ook dat $\mathbb{E}I_i$, de verwachte onbenutte tijd voor patiënt i, klein is. De tijd van de dokter is niet altijd evenveel waard als de tijd die een patiënt moet wachten. Daarom introduceren we een parameter $\omega \in (0,1)$ die de relatieve waarde van de tijd van de dokter representeert. Een mogelijke doelfunctie is dan

$$\omega \sum \mathbb{E}I_i + (1-\omega)\sum \mathbb{E}W_i,$$

waarbij we sommeren over alle patiënten.

Bij deze doelfunctie doet zich nog een probleem voor: $\mathbb{E}I_i$ en $\mathbb{E}W_i$ hangen, naast de planning, ook af van de verdeling van de behandeltijden. Echter, deze verdeling weten we niet, alleen de verwachting en variantie. Omdat we willen dat de planning goed werkt voor alle verdelingen met deze verwachting en variantie, nemen we als doelfunctie

$$\sup\left\{\omega \sum \mathbb{E}I_i + (1-\omega)\sum \mathbb{E}W_i\right\},$$

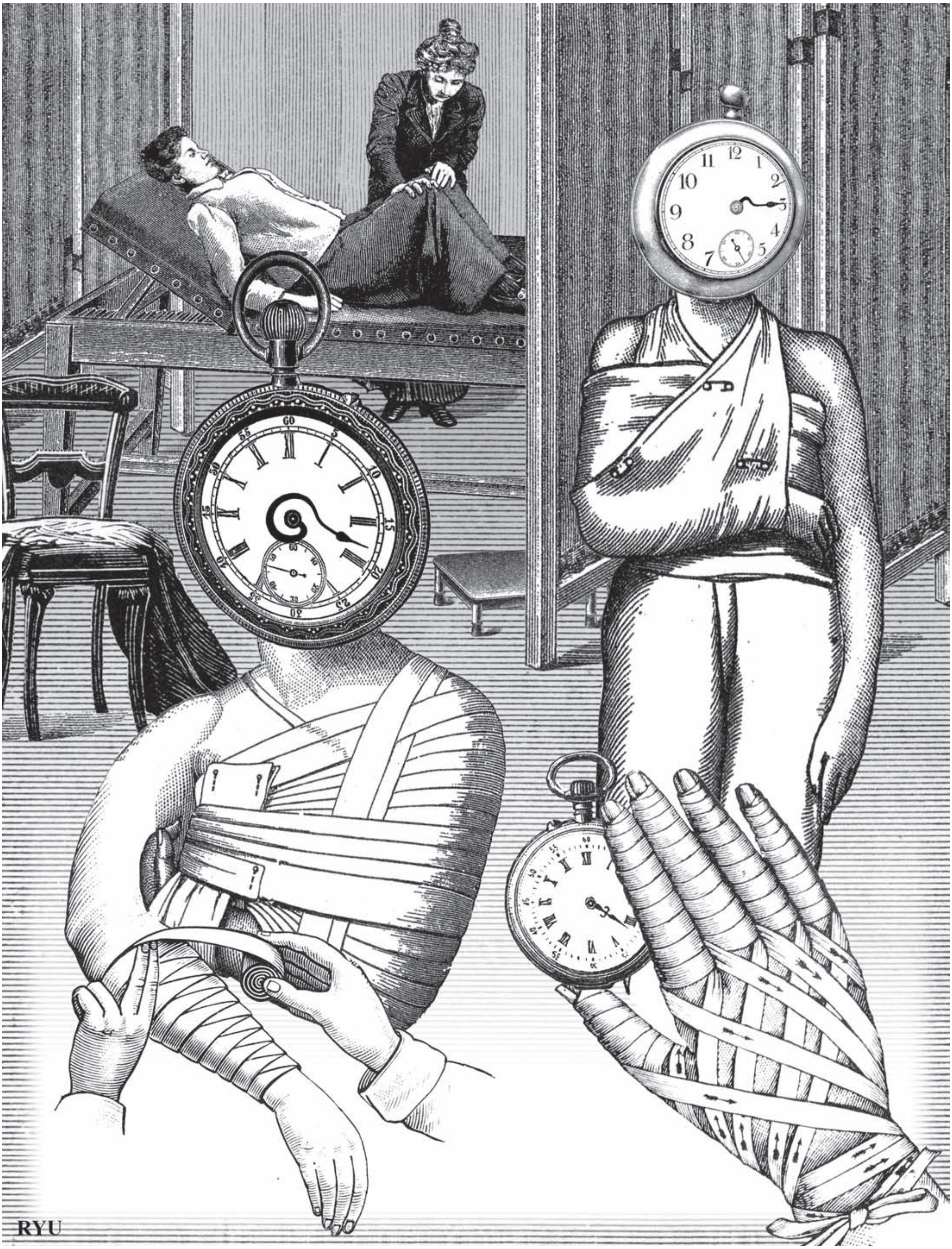
waar we het supremum nemen over alle verdelingen voor de behandeltijden, ieder met de gegeven verwachting en variantie. Deze doelfunctie geeft dan een garantie voor alle mogelijke behandeltijden.

Deze doelfunctie hangt nu nog af van twee zaken: de tijd die per afspraak wordt ingepland, en de volgorde waarin de patiënten behandeld worden. Deze zullen zo gekozen moeten worden dat de doelfunctie minimaal is. In mijn masterscriptie heb ik naar beide deelproblemen onderzoek gedaan, maar niet in precies dezelfde situatie. Het gezamenlijke probleem is dus nog niet opgelost.

Geplande tijd per afspraak

Eerst bekijken we het probleem om te bepalen hoeveel tijd we per afspraak moeten inplannen. Hiervoor nemen we aan dat de behandeltijden van alle patiënten dezelfde verdeling hebben, met gegeven verwachting μ en variantie σ^2. We zullen kijken naar de limiet als het aantal patiënten naar oneindig gaat. Dit geeft ons dan een benadering voor situaties met een eindig maar groot aantal patiënten.

Als er oneindig veel patiënten zijn, zal de situatie van een patiënt niet verschillen van zijn voorganger. We nemen daarom aan dat dezelfde hoeveelheid tijd, die we x zullen noemen, gepland is voor alle afspraken. De verwachte wachttijd, $\mathbb{E}W$, van een



Illustratie: Ryu Tajiri

patiënt is nu hetzelfde voor iedere patiënt. Hetzelfde geldt voor de verwachte ongebruikte tijd tussen twee patiënten, $\mathbb{E}I$. Als doelfunctie krijgen we dan

$$\sup\{\omega\mathbb{E}I + (1-\omega)\,\mathbb{E}W\}.$$

Voor de verwachte ongebruikte tijd is het bekend dat $\mathbb{E}I = x-\mu$, voor iedere verdeling voor de behandeltijd met verwachting μ. Om de doelfunctie te kunnen bepalen moeten we dus nog $\sup\mathbb{E}W$ bepalen. We kunnen nu een bovengrens van Kingman [4] voor algemene wachtrijen toepassen. Dit geeft de volgende bovengrens voor de verwachte wachttijd:

$$\mathbb{E}W \leq \frac{\sigma^2}{2(x-\mu)}.$$

Deze bovengrens is in deze situatie ook strikt (zoals bewezen in [2]), dat wil zeggen er zijn verdelingen voor de behandeltijd met verwachting μ en variantie σ^2 zodanig dat $\mathbb{E}W$ deze bovengrens willekeurig dicht benaderd.

Nu weten we dus dat $\mathbb{E}I = x-\mu$ en $\sup\mathbb{E}W = \frac{\sigma^2}{2(x-\mu)}$. De doelfunctie wordt nu

$$\sup\{\omega\mathbb{E}I + (1-\omega)\,\mathbb{E}W\} = \omega(x-\mu) + (1-\omega)\frac{\sigma^2}{2(x-\mu)}.$$

We zoeken nu de x waarvoor deze doelfunctie minimaal is. Standaard calculus geeft ons dat deze x gegeven wordt door

$$x = \mu + \sigma\sqrt{\frac{1-\omega}{2\omega}}.$$

Volgorde bepalen

Het volgordeprobleem is het probleem om te bepalen in welke volgorde patiënten behandeld moeten worden als er verschillende soorten patiënten zijn. We nemen aan dat we de patiënten kunnen indelen in m groepen van soortgelijke patiënten. Patiënten in dezelfde groep worden achter elkaar in een tijdsblok behandeld. We willen nu weten in welke volgorde de m groepen moeten worden ingedeeld.

We nemen aan dat de behandeltijden van patiënten in dezelfde groep dezelfde verdeling hebben. In groep k heeft de behandeltijd verwachting μ_k en variantie σ_k^2. We plannen μ_k tijd in voor iedere patiënt in groep k, dus de tijd die we inplannen per afspraak is gelijk aan de verwachte tijdsduur van de afspraak (wat niet optimaal hoeft te zijn).

We zullen de limiet bekijken als het aantal patiënten naar oneindig gaat, terwijl de fractie patiënten in groep k, die we T_k noemen, gelijk blijft voor alle k.

We kunnen de wachttijd van een patiënt als volgt uitrekenen: tel bij de wachttijd van de vorige patiënt op hoeveel vertraging de vorige patiënt veroorzaakt of goedmaakt. Als dit positief is, is dit de wachttijd van de huidige patiënt, en als dit negatief is, is de wachttijd van de huidige patiënt nul.

Binnen een groep zijn de vertragingen (de behandeltijd min de geplande tijd) steeds gelijk verdeeld. Wegens de centrale limietstelling zullen sommen van vertragingen van een groot aantal patiënten zich steeds meer volgens een normale verdeling gaan gedragen. Een generalisatie hiervan is de stelling van Donsker, die ons vertelt dat het hele proces van vertragingen voor een groot aantal patiënten zich steeds meer zal gedragen als een zogenaamde 'Brownse beweging'.

Voor de wachttijden zal het proces nu steeds meer gaan lijken op een omhoog geduwde versie van een Brownse beweging. Voor deze omhoog geduwde Brownse beweging kunnen we nu op ieder moment de verwachting berekenen. Deze blijkt voor ieder moment het kleinst te zijn als we de groepen behandelen in volgorde van hun variantie σ_k^2, met de kleinste variantie eerst.

De onbenutte tijd kan ook met behulp van de Brownse beweging benaderd worden. De verwachting van de onbenutte tijd blijkt dan niet van de volgorde af te hangen. In totaal is de volgorde met de kleinste variantie eerst het beste.

Conclusie

Als de behandeltijden van alle patiënten dezelfde verdeling hebben, vonden we dat voor een groot aantal patiënten de optimale behandeltijd gegeven wordt door

$$x = \mu + \sigma\sqrt{\frac{1-\omega}{2\omega}}.$$

We plannen dus voor een afspraak tijd in gelijk aan de verwachting plus een extra tijd lineair aan de standaardafwijking om eerdere vertragingen te kunnen opvangen. Vooral voor ω dicht bij nul, als vooral de wachttijd belangrijk is, is dit opvangen van eerdere vertragingen belangrijk. De formule lijkt dus met die intuïtie overeen te komen.

We keken ook naar het volgordeprobleem, waarbij we te maken hebben met een aantal grote groepen patiënten. Als binnen een groep de behandeltijden dezelfde verdeling hebben, en we plannen voor iedere afspraak tijd in gelijk aan de verwachte behandeltijd, dan is het het beste om de groepen te behandelen in volgorde van de variantie van hun behandeltijd. Ook dit is intuïtief aannemelijk, aangezien groepen met een grote variantie een groot risico vormen om vertragingen te veroorzaken, en dus het beste als laatste behandeld kunnen worden.

De twee deelresultaten sluiten helaas nog niet helemaal op elkaar aan. Bij het volgordeprobleem namen we aan dat de geplande tijd per afspraak gelijk is aan de verwachte behandeltijd, terwijl we eerder al zagen dat we eigenlijk meer tijd zouden moeten inplannen. Het gezamenlijke probleem van zowel de optimale behandeltijden en de optimale volgorde is dus nog steeds onopgelost.

Referenties

1. A. Ahmadi-Javid, Z. Jalali en K. Klassen, Outpatient appointment systems in healthcare: a review of optimization studies, *European Journal of Operational Research* 258(1) (2017), 3–34.
2. D. Daley, A. Kreinin en C. Trengove, Inequalities concerning the waiting-time in single-server queues: a survey, in: *Queueing and Related Models*, Clarendon Press, 1992, pp. 177–223.
3. M.A. de Kemp, *Robust Appointment Scheduling and Sequencing for Many Patients*, M.Sc. Thesis, VU Amsterdam, 2016.
4. J.F.C. Kingman, Some inequalities for the GI/G/1 queue, *Biometrika* 49(3-4) (1962), 315–324.