Eri vaihtoehtojen tutkiminen haarautuvassa monivaiheisessa näytteenottoketjussa

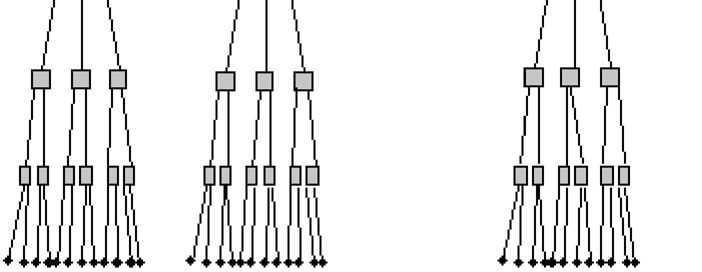
1 Näytteenotto ilman optimointia kivihiilen tuhkapitoisuuden määrittämiseksi

Kivihiilitoimitus, jonka kokonaismäärä on 1000 tonnia, koostuu 20 rautatievaunusta, joiden koko on 50 t. Lastin purkamisen yhteydessä otetaan joka toisesta vaunusta 3 primäärinäytettä, joiden koko on 1.6 kg. Näistä jokainen jauhetaan erikseen ja jaetaan jakolaitteella 16 osanäytteeksi à 100 g. Näistä kaksi lähetetään laboratorioon tutkittavaksi. Laboratoriossa kustakin näytteestä valmistetaan kaksi 1 g:n suuruista analyysinäytettä, jotka poltetaan tuhkapitoisuuden määrittämiseksi punnitsemalla. Esikokeiden perusteella tiedetään, että vaunujen välinen keskihajonta tuhkan massaosuutena on 0.045, vaunun sisältä otetun 1.6 kg:n suuruisen primäärinäytteen keskihajonta on 0.022, primäärinäytteistä valmistettujen laboratorionäytteiden välinen keskihajonta on 0.0030 ja samasta laboratorionäytteestä valmistettujen analyysinäytteiden keskihajonta on 0.0040.

Tässä näytteenottoketjussa on neljä tasoa. Määritellään aluksi näytteenottotasojen suhteelliset koot. Koko voidaan määritellä käyttämällä mittana kultakin tasolta otettujen näytteiden kokoa:

Taso 1	$N_1 := 20$			$n_1 := 10$
Taso 2	$N_2 := \dfrac{50 \cdot 10^3 \cdot kg}{1.6 \cdot kg}$	==>	$N_2 = 31250$	$n_2 := 3$
Taso 3	$N_3 := \dfrac{1600 \cdot g}{100 \cdot g}$	==>	$N_3 = 16$	$n_3 := 2$
Taso 4	$N_2 := \dfrac{100 \cdot g}{1.0 \cdot g}$	==>	$N_2 = 100$	$n_4 := 2$

Määritysketju voidaan siis kuvata seuraavalla kaaviolla:



N_1=20, n_1 =10, $s_1^2 = 2 \cdot 10^{-3}$

N_2=31250, n_2 =3, $s_2^2 = 5 \cdot 10^{-4}$

N_3=16, n_3 =2, $s_3^2 = 9 \cdot 10^{-6}$

N_4=100, n_4 =2, $s_4^2 = 1.6 \cdot 10^{-7}$

Analysoitavien näytteiden luku, $n = n_1 \cdot n_2 \cdot n_3 \cdot n_4 = 120$

Ylintä vaihetta lukuun ottamatta voidaan olettaa, että $N_i >> n_i$, joten

$$s_x^2 = \frac{N_1 - n_1}{N_1 - 1} \cdot \frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_1 \cdot n_2} + \frac{(s_3)^2}{n_1 \cdot n_2 \cdot n_3} + \frac{(s_4)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4}$$

$$= \frac{20-10}{20-1} \cdot \frac{2 \cdot 10^{-3}}{10} + \frac{5 \cdot 10^{-4}}{10 \cdot 3} + \frac{9 \cdot 10^{-6}}{10 \cdot 3 \cdot 2} + \frac{1.6 \cdot 10^{-7}}{10 \cdot 3 \cdot 2 \cdot 2}$$

$$= 1.053 \cdot 10^{-4} + 1.667 \cdot 10^{-5} + 1.50 \cdot 10^{-7} + 1.33 \cdot 10^{-9} = 1.22 \times 10^{-4}$$

Erän keskiarvon kokonaiskeskihajonta, $s_x = 0.011$

ja kokonaisepävarmuus (95 %:n luottamustasolla), $U_{.95} = 2 \cdot s_x = 0.022$
==============

2 Kustannusten suhteen optimoitu otanta

Jotta määritysketju voitaisiin optimoida kustannusten suhteen, tarvitaan eri työvaiheiden yksikkökustannukset, joko suhteellisina tai absoluuttisina. Oletetaan, että tässä pätevät seuraavat kustannussuhteet.
Vaihe 1 koostuu pelkästään vaunun valinnasta, joten $c_1 = 0$. Pidetään perustana yhden tuhkanmäärityksen hintaa. Koska analyysinäytteiden punnitus ja poltto voidaan tehdä sarjassa, jää yksikkökustannus pieneksi, $c_4 = c'$. Toinen ja kolmas vaihe vaativat näytettä kohti enemmän työtä, koska ne sisältävät näytteen jauhatuksen ja näytteen jaon vaatiman työn. Tässä tapauksessa $c_2 = c_4 = 10 \cdot c'$. A .kohdan kokonaiskulut ovat siten:

$$c := n_1 \cdot c_1 + (n_1 \cdot n_2) \cdot c_2 + (n_1 \cdot n_2 \cdot n_3) \cdot c_3 + (n_1 \cdot n_2 \cdot n_3 \cdot n_4) \cdot c_4$$
$$= 10 \cdot 0 + (10 \cdot 3) \cdot 10c' + (10 \cdot 3 \cdot 2) \cdot '10 \cdot c' + (10 \cdot 3 \cdot 2 \cdot 2) \cdot c'$$
$$= 0 + 300 \cdot c' + 600 \cdot c' + 120 \cdot c' = 1020c'$$

Minimoidaan kokonaisvarianssi siten, että kokonaiskustannukset eivät saa ylittää A.-kohdan kuluja eli

$$c_t = 1020 c'$$

Koska suurin osa kokonaisvarianssista johtuu vaunun valinnasta, eikä vaunun valinta aiheuta kustannuksia, niin tämän vaiheen varianssi kannattaa eliminoida. Kokonaisvarianssin yhtälöstä nähdään, että ottamalla näyte jokaisesta vaunusta eliminoituu vaunujen välisen varianssin vaikutus kokonaan keskiarvon varianssista. Jos siis valitaan $n_1 := N_1 = 20$, ja sijoitetaan tämä kokonaisvarianssin yhtälöön, niin nähdään, että optimoitavaksi jää kolmivaiheinen määritysketju:

$$s_x^2 = \frac{1}{n_1}\cdot\left[\frac{(s_2)^2}{n_2} + \frac{(s_3)^2}{n_2\cdot n_3} + \frac{(s_4)^2}{n_2\cdot n_3\cdot n_4}\right]$$

Optimoidun otannan kaavat antavat neljännen ja kolmannen vaiheen näytteiden luvuksi seuraavat tulokset:

$$n_4 := \frac{s_4}{s_3}\cdot\sqrt{\frac{c_3}{c_4}} = \frac{4\cdot 10^{-4}}{3\cdot 10^{-3}}\cdot\sqrt{\frac{10\cdot c'}{c'}} = 0.422 \qquad ==> \quad \text{valitaan } n_4 := 1$$

$$n_3 := \frac{s_3}{s_2}\cdot\sqrt{\frac{c_2}{c_3}} = \frac{3\cdot 10^{-3}}{2.2\cdot 10^{-2}}\cdot\sqrt{\frac{10\cdot c'}{10\cdot c'}} = 0.136 \qquad ==> \quad \text{valitaan } n_3 := 1$$

Vaunuista otettavien primäärinäytteiden luku voidaan nyt ratkaista kokonaiskulujen yhtälöstä, koska siitä tunnetaan kaikki muut tekijät n_2:ta lukuun ottamatta.

$$c_t = 1020\,c' = (n_1\cdot n_2)\cdot 10\cdot c' + (n_1\cdot n_2\cdot n_3)\cdot 10\cdot c' + (n_1\cdot n_2\cdot n_3\cdot n_4)\cdot c'$$

$$= (20\cdot n_2)\cdot 10\cdot c' + (20\cdot n_2\cdot 1)\cdot 10\cdot c' + (20\cdot n_2\cdot 1\cdot 1)\cdot c'$$

$$==> \quad n_2 := \frac{1020}{420} = 2.43 \qquad ==> \quad \text{valitaan } n_2 := 2$$

(Koska kokonaiskulut oli määrätty, niin pyöristys on tehtävä aina alaspäin)

Ketjun kokonaisvarianssiksi siis saadaan:

$$s_x^2 = \frac{N_1 - n_1}{N_1 - 1}\cdot\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_1\cdot n_2} + \frac{(s_3)^2}{n_1\cdot n_2\cdot n_3} + \frac{(s_4)^2}{n_1\cdot n_2\cdot n_3\cdot n_4}$$

$$= \frac{20-20}{20-1}\cdot\frac{2\cdot 10^{-3}}{20} + \frac{5\cdot 10^{-4}}{20\cdot 2} + \frac{9\cdot 10^{-6}}{20\cdot 2\cdot 1} + \frac{1.6\cdot 10^{-7}}{20\cdot 2\cdot 1\cdot 1}$$

$$= 0 + 1.237\cdot 10^{-5} + 2.25\cdot 10^{-7} + 4.0\cdot 10^{-9} = 1.26\times 10^{-5}$$

Erän keskiarvon kokonaiskeskihajonta,

$$s_x = 3.57\times 10^{-3}$$

ja kokonaisepävarmuus (95 %:n luottamustasolla),

$$U_{.95} = 2\cdot s_x = 7.1\times 10^{-3}$$

==============

Määritysketjun kokonaiskulut ovat: $c = (20 \cdot 2) \cdot 10 \cdot c' + (20 \cdot 2) \cdot 10 \cdot c' + (20 \cdot 2) \cdot c' = 840c'$

Tässä tapauksessa optimoinnilla saavutetaan huomattava hyöty: Analyysikustannukset alenevat n. 20 % ja keskiarvon epävarmuus pienenee n. kolmasosaan edelliseen määritysketjuun verrattuna.

3 Kustannusten minimointi, kun kokonaisepävarmuus on määrätty

Optimoi näytteenottosuunnitelma, kun tavoitteeksi asetetaan, että erän tuhkapitoisuus-määrityksen kokonaisepävarmuus saa olla enintään 0.005.

$$U_{.95} := 0.005 = 2 \cdot s_x \quad ==> \quad \text{tavoite} \quad s_T := \frac{0.005}{2} \quad \text{ja} \quad {s_T}^2 = 6.25 \times 10^{-6}$$

Tasojen 1,3 ja 4 osalta pätevät edelleen samat päättelyt, kuin edellisessä kohdassa, joten

$$n_1 = 20 \quad , \quad n_3 = 1 \quad \text{ja} \quad n_4 = 1$$

n_2 on nyt ratkaistava kokonaisvarianssin yhtälöstä, johon nämä arvot sijoitetaan:

$${s_T}^2 = \frac{N_1 - n_1}{N_1 - 1} \cdot \frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_1 \cdot n_2} + \frac{(s_3)^2}{n_1 \cdot n_2 \cdot n_3} + \frac{(s_4)^2}{n_1 \cdot n_2 \cdot n_3 \cdot n_4}$$

$$==> \quad n_2 := \frac{(s_2)^2 + (s_3)^2 + (s_4)^2}{n_1 \cdot {s_T}^2} = \frac{5 \cdot 10^{-4} + 9 \cdot 10^{-6} + 1.6 \cdot 10^{-7}}{20 \cdot 6.25 \cdot 10^{-6}} = 4.073$$

Valitaan $n_2 := 4$

Oikeastaan pyöristys pitäisi suorittaa ylöspäin, koska tavoitevarianssi on määrätty, mutta tässä tapauksessa pyöristys aiheuttaa niin pienen ylityksen, ettei sillä käytännössä ole merkitystä, (varsinkin, kun approksimaatio N >>n yliarvioi hieman varianssin). Tarkistetaan kuitenkin tulos:

$${s_x}^2 = \frac{(s_2)^2 + (s_3)^2 + (s_4)^2}{n_1 \cdot n_2} = \frac{5 \cdot 10^{-4} + 9 \cdot 10^{-6} + 1.6 \cdot 10^{-7}}{20 \cdot 4} = 6.365 \times 10^{-6}$$

$$s_x = 2.52 \times 10^{-3}$$

Määritysketjun kokonaiskulut ovat nyt:

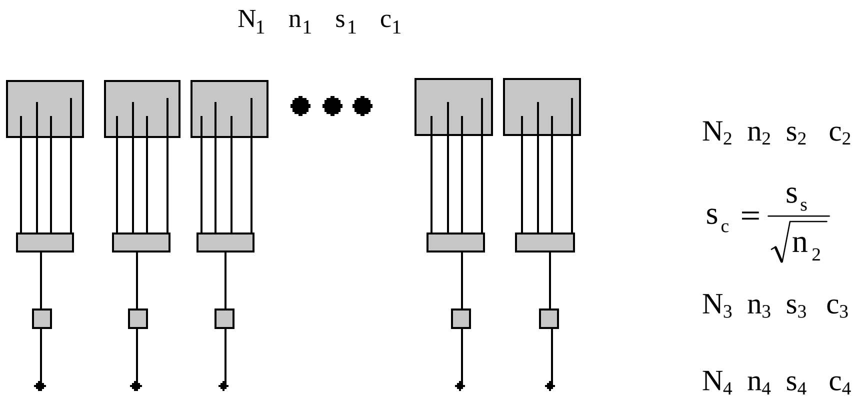
$$c = (20 \cdot 4) \cdot 10 \cdot c' + (20 \cdot 4) \cdot 10 \cdot c' + (20 \cdot 4) \cdot c' = 1680c'$$

Kuten havaitaan vertaamalla B- ja C-kohdan kustannuksia, niin keskihajonnan pienentäminen 0.00355:stä 0.0025:een, kun määritysketju oli jo optimoitu, kaksinkertaisti kustannukset.

Kun määritysketju on optimoitu, niin tavoitteet kannattaa asettaa realistisesti sellaiselle tasolle, mitä tulosten käyttötarkoitus edellyttää. Jos ne asetetaan liian tiukoiksi, niin siitä joutuu maksamaan tarpeettoman suuren hinnan.

4 Analyysikulujen pienentäminen yhdistämällä näytteitä koonta- näytteeksi

Kohdissa C ja D esitetty analyysikaavio ei vielä välttämättä ole kustannuksiltaan edullisin. Tarkastelemalla tuloksia, havaitaan että prosessin suurin virhelähde on edelleen primäärinäytteen otto vaunusta. Tämän vaikutusta voidaan pienentää ainoastaan lisäämällä näytteiden lukumäärää. Noudatettaessa edellä kuvattua haarautuvaa määrityskaaviota kasvavat kuitenkin analyysinäytteiden luku ja myös kustannukset nopeasti. Määritysten lukua voidaan kuitenkin pienentää yhdistämällä esim. samasta vaunusta otetut näytteet. Nämä näytteet jauhetaan ja homogenisoidaan huolellisesti ja jatketaan siitä eteenpäin valmistamalla kustakin laboratorio- ja analyysinäytteet alla esitetyn kaavion mukaisesti. Luonnollisesti koontanäytteen homogenisointi ja jako nostaa laboratorionäytteen valmistuskustannuksia, koska tällöin joudutaan käsittelemään isompaa näytemassaa.



n_2 on yhdistettävien primäärinäytteiden luku ja c_3 yhden laboratorionäytteen valmistuskulut koontanäytteestä, joka sisältää koontanäytteen homogenisointi- ja jakokulut. Koontanäytteen varianssi pienenee osanäytteiden luvun lisääntyessä samoin kuin keskiarvon varianssi määritysten luvun lisääntyessä.

Koontanäytteen varianssi, s_c on siten:

$s_c^2 := \frac{(s_2)^2}{n_2}$. Muuten merkinnät ovat samat kuin edellisissä kohdissa

Tämän ketjun kokonaisvarianssi, s_x^2 on

$$s_x^2 = \frac{N_1 - n_1}{N_1 - 1} \cdot \frac{(s_1)^2}{n_1} + \frac{N_2 - n_2}{N_2 - 1} \cdot \frac{(s_2)^2}{n_1 n_2} + \frac{N_3 - n_3}{N_3 - 1} \cdot \frac{(s_3)^2}{n_1 \cdot n_3} + \frac{N_4 - n_4}{N_4 - 1} \cdot \frac{(s_4)^2}{n_1 \cdot n_3 \cdot n_4}$$

ja kokonaiskulut

$$c := n_1 \cdot c_1 + (n_1 \cdot n_2) \cdot c_2 + (n_1 \cdot n_3) \cdot c_3 + (n_1 \cdot n_3 \cdot n_4) \cdot c_4$$

Jos käytetään alla annettuja arvoja (koontanäytteen homogenisointi ja jako kaksinkertaistaa laboratorio- näytteen valmistuskulut):

$n_1 = 20$ $n_2 = 4$ $n_3 = 1$ ja $n_4 = 1$ ja $N_i >> n_i$

$c_1 = 0$ $c_2 = 10c'$ $c_3 := 20 \cdot c'$ ja $c_4 = 1c'$

niin kokonaiskeskihajonnaksi saadaan

$$s_x^2 = \left[\frac{(s_2)^2}{n_1 \cdot n_2} + \frac{(s_3)^2}{n_1} + \frac{(s_4)^2}{n_1}\right] = 6.71 \times 10^{-6} \quad \text{ja} \quad s_x = 2.59 \times 10^{-3}$$

==========

Ja kokonaiskuluiksi

$$c = (n_1 \cdot n_2) \cdot 10 \cdot c' + (n_1) \cdot 20 \cdot c' + (n_1) \cdot c' = 1220c'$$

=====

Kustannukset putosivat tässä vaihtoehdossa huomattavasti keskihajonnan kasvaessa vain hieman. Lasketaan vielä tulokset, jos primäärinäytteitä otetaan 5 tai 6 kappaletta vaunua kohti.

Arvolla $n_2 := 5$ $s_x^2 = 5.46 \times 10^{-6}$ $s_x = 2.34 \times 10^{-3}$ ja $c = 1420c'$

$n_2 := 6$ $s_x^2 = 4.62 \times 10^{-6}$ $s_x = 2.15 \times 10^{-3}$ ja $c = 1620c'$

Näistä viimeinenkin vaihtoehto on vielä halvempi kuin c)-kohdan ratkaisu, mutta kokonais- keskihajonta on silti pienempi.

Jos näytteitä yhdistetään niin tuloksista ei voida arvioida, onko vaunujen sisäinen keskihajonta se, mitä laskuissa on käytetty. Jos tämä tieto pitää estimoida tai tarkistaa, niin silloin osanäytteitä ei luonnollisestikaan saa yhdistää.