

Matematiikan kirjoitusohjeita tilastotieteilijöille

Kirjoita suomea matemaattisten symbolien sijasta, jollei se vie kohtuuttomasti tilaa. Numeroita ja symboleita ei mielellään kirjoiteta peräkkäin (ei "lauseessa 3 x on"). Virkkeen aloittaminen numerolla tai symbolilla on mahdollista, mutta monet lukijat kaihtanevat sitä. Tällaista aloitusta ei pidä tehdä, jos edellinen virke päättyy numeroon tai symboliin.

Kaavat ja suhteellisen ilmeisetkin merkinnät selitetään sanoin. Varaudu merkinnöissä etukäteen esimerkiksi mallin laajennukseen. Kaava kirjoitetaan omalle rivilleen, jos se on tärkeä, jos se olisi epäselvä tekstin seassa tai jos siinä esitetään uusi tärkeä merkintä. Vain tärkeät kaavat numeroidaan. Kaavaan viitattaessa lisätään numeron eteen sopiva määre ("yhtälössä (5)"). Kaavaan, lausekkeeseen tms. matemaattiseen merkintään viitataan pronomiinilla joka. Pitkä kaava tekstistä erotettuna katkaistaan ennen yhteenlasku-, kerto- tms. merkkiä mutta tekstin seassa niiden jälkeen.

Käännä käsitteet suomeksi, ellei vierasperäinen ilmaus ole juurtunut tilastotieteeseen (esim. lause eikä teoreema ja aineisto eikä data). Merkitse aineiston muuttujat mahdollisimman yksinkertaisesti ja ilmeisesti (kulutusta c :llä, hintaa p :llä jne.).

Pienet luvut ($1, \dots, 20$) ja tasaluvut (esim. 100 tai 1000) kirjoitetaan tavallisesti kirjaimilla. Luvut merkitään numeroilla, jos halutaan ilmaista täsmällinen määrä (kuten esimerkiksi harjoitustehtävissä tahdotaan), tekstissä on useita lukuja, joita vertaillaan toisiinsa tai lukua käytetään nimen tapaan ("lauseessa 3 osoitetaan"). Suomen kielessä desimaaliluvun merkki on pilkku eikä piste: 3,141 eikä 3.141. Desimaaliluvuissa riittää lopputuloksissa usein yksi tai kaksi desimaalia (mieti, mikä on järkevä raportointitarkkuus).

Muuttujat ja parametrit erotetaan pienillä väleillä yhteenlasku-, kerto- yms. merkeistä ja kursivoidaan: $x > y$ eikä $x > y$. Funktion merkki erotetaan pienellä välillä funktion argumentista, ellei sitä rajata sulkeilla: $\log 10$ mutta $\log(10)$. Yleisimpien matemaattisten funktioiden merkkiä (\sin , \lim jne.) ei kursivoida. (L^AT_EX huomioi edelliset asiat automaattisesti matematiikka-asetuksen ollessa päällä.)

Ellipsis-merkki (\dots) on luetelmissa ja tuloissa alhaalla ($1, \dots, T$ ja $a_1 \dots a_T$) mutta muulloin yleensä keskikorkeudella ($a_1 + \dots + a_T$). Huomaa, että $n \times k$ -matriisissa on väli mutta $x - y$ -akselistossa ei. Vektorit ja matriisit kirjoitetaan lihavoiduilla pienillä ja isoilla kirjaimilla: \mathbf{y} ja \mathbf{X} .

Erityisesti tilastotieteellisiin laskutoimituksiin liittyvät tavallisimmat merkinnät kirjoitetaan Sans serif -kirjasinlajeilla (esim. odotusarvo, varianssi, todennäköisyys ja selitysosuus):

$$E(Y), V(Z), P(A) \text{ ja } R^2.$$

Jakaumat osoitetaan niin ikään Sans serif -kirjasinlajeilla (esim. normaali-, binomi- ja tasainen jakauma):

$$N(\mu, \sigma^2), \text{Bin}(n, p) \text{ ja } U(a, b).$$

Satunnaismuuttujat kirjoitetaan isoilla ja niiden realisaatiot pienillä kirjaimilla:

$$P(Y = 3) = 1/6$$

ja

$$y = 3.$$

Osamäärissä tulee huolehtia, että ne tulostuvat riittävän isoina: Ei

$$\frac{a}{b} = \frac{c}{d}$$

vaan

$$\frac{a}{b} = \frac{c}{d}.$$

Sama kaava on usein hyvä kirjoittaa erilailla tekstin seassa kuin tekstistä erotettuna: $a/b = c/d$.

Yhtälöt ja muu matematiikka voidaan sijoittaa tekstiin, tai erilleen tekstistä, mutta pilkutetaan aina muiden lauseiden ja lauseenjäsenien tavoin. Yhtälöön

$$\begin{aligned} y(x) &= \frac{\alpha}{2} + \beta x \\ &= \alpha^* + \beta x \end{aligned}$$

ei tule pilkkuja, mutta kahden eri yhtälön väliin tulee pilkku:

$$\begin{aligned} y(x) &= \frac{\alpha}{2} + \beta x, \\ z(x) &= \gamma + \delta x. \end{aligned}$$

Matematiikka pilkutetaan tarpeen mukaan:

$$y(x) = \begin{cases} 1, & \text{jos } x \geq 0, \\ 0, & \text{jos } x < 0. \end{cases}$$

Estimoidut yhtälöt voi raportoida esimerkiksi näin:

$$y = \begin{matrix} 8,430 & + & 0,114x_1 & - & 0,001x_2 & + & 0,169x_3 & + & 0,339x_4 & + & \hat{\varepsilon} \\ (0,000) & & (0,840) & & (0,983) & & (0,747) & & (0,000) & & \end{matrix}$$

$$R^2 = 0,256, F_{4,132} = 11,269, n = 137.$$

Ohjeiden tärkein lähde on N. Higham (1998): *Handbook of Writing for the Mathematical Sciences. 2. laitos*. SIAM, Philadelphia. Lisäneuvoja voi hakea artikkelista K. M. Abadir ja J. R. Magnus (2002): Notation in Econometrics: A Proposal for a Standard. *Econometrics Journal*, 5, 76–90. Edellä on paikoin poikettu lähteiden ehdotuksista. Kiitän Pekka Niemistä erinomaisista huomioista.

23.1.2013,

yliopistonlehtori Pekka Pere
aikasarja-analyysin ja ekonometrian linja
sosiaalitieteiden laitos
Helsingin yliopisto

P.S. \LaTeX mahdollistaa kauniin matemaattisen esityksen lisäksi yksityiskoh-
taisen taulukon ominaisuuksien määrittelyn. Katso tämän tekstin tuottaneesta
 \LaTeX -tiedostosta, miten tämä taulukko

seksikump- panien lkm	onnellisuus			yhteensä (%)
	ei kovin	melko	hyvin	
0	18,82	55,29	25,88	100
1	7,95	56,03	36,03	100
≥ 2	18,27	63,46	18,27	100
yhteensä	12,00	56,81	31,19	100

on tehty. (Tämän taulukon typografiaa voi vielä parantaa, mutta se on toinen
tarina.)