Johdatus tilastolliseen päättelyyn, 1. harjoitus (Viikko 12: 19.–22.3.2013)

1. Parametrilla θ on tässä tehtävässä kaksi vaihtoehtoista arvoa: 1 tai 2. Diskreetin satunnaismuuttujan Y pistetodennäköisyysfunktion $g(y;\theta)$ arvot on taulukoitu alla. (Tavalliseen tapaan sovimme, että $g(y;\theta) = 0$ kaikilla niillä arvoilla y, joita ei ole mainittu.)

y	1	2	3	4	5
$g(y;1)$	0	0.3	0.4	0.2	0.1
$g(y;2)$	0.1	0.4	0.3	0.1	0.1

Tarkista, että kumpikin funktioista $g(y;1)$ ja $g(y;2)$ on pistetodennäköisyysfunktio. Esitä kumpikin funktio graafisesti.

2. (Jatkoa edelliseen tehtävään.) Emme tiedä, kumpi parametrinarvo on oikea. Suoritamme satunnaiskokeen, jonka tuloksena havaitaan satunnaismuuttujan Y arvo y.

a) Tulos oli $y = 1$. Tässä tilanteessa voimme sanoa varmuudella, kumpi parametrinarvo on oikea. Miksi?

b) Tulos oli $y = 2$. Kummalla parametrinarvolla tämä havainto on todennäköisempi?

3. Merkintä $\sum_{i=1}^{n} x_i$ tarkoittaa reaalilukujen $x_1, \dots, x_n$ summaa, eli

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n.$$

Summamerkintää käytetään usein sekaisin aritmeettisten operaatioiden kanssa. Tottumattoman voi olla vaikea hahmottaa, mihin asti summaoperaattorin vaikutus kaavassa ulottuu. (Merkintöjä sujuvasti käyttävän henkilön voi toisaalta olla vaikea ymmärtää, että kukaan voisi tälläisiä merkintöjä ymmärtää väärin.) Nyt harjoittelemme summamerkintöjen käyttöä.

Osa seuraavista kaavoista on identiteettejä (eli ne ovat voimassa kaikilla järkevillä kaavoissa esiintyvien suureiden n, a, b, x_1, …, x_n ja $y_1, \dots, y_n$ arvoilla). Joku tai jotkut kaavoista eivät ole identiteettejä (vaikka ne saattavat päteä joillakin suureiden arvoilla). Aukaise summamerkinnät (kolmen pisteen avulla), ja tarkista tavanomaisien aritmetiikan laskusääntöjen avulla, onko annettu kaava identiteetti. Jos se ei ole identiteetti, korjaa kaavan oikea puoli niin että saat aikaan identiteetin.

$$\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i \tag{1}$$

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i \tag{2}$$

$$\sum_{i=1}^{n} (a x_i + b) = a \sum_{i=1}^{n} x_i + b. \tag{3}$$

Huomautuksia: Kaavan (1) vasemmalla puolella lasketaan yhteen luvut $ax_1, \dots, ax_n$. Kaavan (2) vasemmalla puolella lasketaan yhteen luvut $x_1 + y_1, \dots, x_n + y_n$ ja kaavan oikealla puolella lasketaan yhteen summat $s_x = \sum x_i$ ja $s_y = \sum y_i$. Kaavan (3) vasemmalla puolella summataan luvut $ax_1 + b, \dots, ax_n + b$ ja kaavan oikealla puolella lasketaan yhteen tulo as_x sekä luku b.

4. Olkoot $y_1, \ldots, y_n$ reaalilukuja, ja olkoon $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ niiden aritmeettinen keskiarvo. Näytä, että mille tahansa $a \in \mathbb{R}$ pätee kaava

$$\sum_{i=1}^n (y_i - a)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - a)^2.$$

Opastus: lähde liikkeelle esityksestä $(y_i - a)^2 = (y_i - \bar{y} + \bar{y} - a)^2$.

5. Olkoon X satunnaismuuttuja ja a reaaliluku. Näytä, että

$$E(X - a)^2 = \operatorname{var} X + (\mu - a)^2.$$

Tässä $\mu = EX$ on X:n odotusarvo ja $\operatorname{var} X = E(X - \mu)^2$ on X:n varianssi.
Opastus: lähde liikkeelle esityksestä $(X - a)^2 = (X - \mu + \mu - a)^2$.

6. Merkintä $\prod_{i=1}^n x_i$ tarkoittaa reaalilukujen $x_1, \ldots, x_n$ tuloa, eli

$$\prod_{i=1}^n x_i = x_1 x_2 \cdots x_n.$$

Osa seuraavista kaavoista on identiteettejä ja osa taas ei. Aukaise tulomerkinnät (kolmen pisteen avulla) ja tarkista kustakin kaavasta, onko se identiteetti. Korjaa epäidenttisten kaavojen oikea puoli niin, että saat aikaan identiteetin, mikäli tämä on helposti tehtävissä. (Vaikeammissa tapauksissa voit tyytyä antamaan vastaesimerkin.)

$$\left(\prod_{i=1}^n x_i\right)\left(\prod_{i=1}^n y_i\right) = \prod_{i=1}^n x_i y_i \tag{4}$$

$$\prod_{i=1}^n x_i + \prod_{i=1}^n y_i = \prod_{i=1}^n (x_i + y_i) \tag{5}$$

$$\prod_{i=1}^n a x_i = a \prod_{i=1}^n x_i \tag{6}$$

$$\log(\prod_{i=1}^n x_i) = \sum_{i=1}^n \log(x_i), \qquad \text{kun } x_i > 0 \text{ kaikilla } i. \tag{7}$$

$$\exp(\sum_{i=1}^n y_i) = \sum_{i=1}^n \exp(y_i). \tag{8}$$

Huomautus: Tällä kurssilla $\log(x)$ tarkoittaa luonnollista logaritmia, eli $\log(x) \equiv \ln(x)$. Merkintä $\exp(x)$ tarkoittaa samaa kuin e^x.