Tehtävä 4

Olkoon $f : (0, \infty) \to \mathbb{R}$ jatkuva ja vähenevä funktio, jolle pätee

$$f(xy) = f(x) + f(y)$$

kaikilla $x, y \in (0, \infty)$.

(a) Osoita, että $f(1) = 0$.

(b) Oletetaan, että $f(4) = A$. Määritä $f(2)$ sekä $f(8)$.

(c) Osoita, että f on välttämättä muotoa $f(x) = -\log_r(x)$ jollakin $r > 0$.

Ratkaisu:

(a) Oletuksesta $f(1) = f(1 \cdot 1) = f(1) + f(1) = 2f(1)$. Jos $f(1) \neq 0$, tällöin saadaan ristiriita $1 = 2$ eli täytyy olla $f(1) = 0$.

(b) Koska $f(4) = A$, pätee $A = f(4) = f(2 \cdot 2) = f(2) + f(2) = 2f(2)$, josta saadaan $f(2) = A/2$. Vastaavasti $f(8) = f(2 \cdot 4) = f(2) + f(4) = A/2 + A = 3A/2$.

(c) Osoitetaan aluksi, että $f(x) = -\log_r(x)$ toteuttaa annetut ehdot. Selvästi kyseessä on jatkuva funktio ja koska

$$f'(x) = -\frac{1}{x \log(r)} < 0$$

kaikilla $x > 0$, kyseessä on myös vähenevä funktio. Lisäksi logaritmin laskusäännöillä

$$f(xy) = -\log_r(xy) = -\log_r(x) - \log_r(y) = f(x) + f(y)$$

eli f toteuttaa kaikki annetut ehdot.

Toisen suunnan (vain $f(x) = -\log_r(x)$ kelpaa ratkaisuksi) osoittamiseen todistetaan ensin, että Cauchyn funktionaaliyhtälön

$$g(x + y) = g(x) + g(y)$$

ratkaisee vain ja ainoastaan $g(x) = \lambda x$ jollakin $\lambda \in \mathbb{R}$.

Oletetaan siis, että jatkuva funktio $g : \mathbb{R} \to \mathbb{R}$ toteuttaa Cauchyn funktionaaliyhtälön. Tällöin $g(0) = g(0 + 0) = g(0) + g(0) = 2g(0)$, joten $g(0) = 0$. Toisaalta kaikilla $x \in \mathbb{R}$

$$g(x) + g(-x) = g(x + (-x)) = g(0) = 0 \iff g(-x) = -g(x).$$

Induktiivisesti nähdään, että $g(nx) = ng(x)$ kaikilla $n \in \mathbb{N}$. Lisäksi g:n parittomuuden perusteella tämä voidaan yleistää kaikille $n \in \mathbb{Z} \setminus \{0\}$, sillä jos $n < 0$, pätee

$$g(nx) = -g(-nx) = -(-ng(x)) = ng(x).$$

Jos taas $n = 0$, pätee triviaalisti $g(nx) = g(0) = 0 = 0g(x)$. Nyt siis erityisesti $g(n) = ng(1)$ kaikilla $n \in \mathbb{Z}$. Olkoon sitten $m, n \in \mathbb{Z}$ siten, että $n \neq 0$. Tällöin

$$g(m) = g\left(n \ \cdot \frac{m}{n}\right) = mg(1).$$

Toisaalta

$$g(m) = g\left(n \cdot \frac{m}{n}\right) = ng(m/n),$$

joten

$$g(m/n) = \frac{g(m)}{n} = \frac{mg(1)}{n}.$$

Siispä kaikilla $x \in \mathbb{Q}$ pätee $g(x) = xg(1)$. Koska g on jatkuva, reaalianalyysin tulosten perusteella $g(x) = xg(1)$ pätee myös kaikilla $x \in \mathbb{R}$. Löydettiin siis $\lambda = g(1)$ siten, että $g(x) = \lambda x$ ja Cauchyn funktionaaliyhtälöllä on täsmälleen tämä ratkaisu, sillä kaikki ratkaisuvaiheet ovat yhtäpitäviä.

Asetetaan nyt $h(x) = f(e^x)$, missä f toteuttaa alkuperäiset ehdot. Koska h on jatkuva ja

$$h(x + y) = f(e^{x+y}) = f(e^x e^y) = f(e^x) + f(e^y) = h(x) + h(y),$$

h toteuttaa Cauchyn funktionaaliyhtälön ja siten välttämättä $h(x) = f(e^x) = \lambda x$ jollakin $\lambda \in \mathbb{R}$. Tällöin $f(x) = h(\log(x)) = \lambda \log(x)$. Toisaalta

$$\log_r(x) = \frac{\log(x)}{\log r},$$

joten huomioimalla vähenevän funktion vaatimuksen saadaan $\lambda = -1/\log r$ joten f on välttämättä muotoa

$$f(x) = \lambda x = -\frac{\log(x)}{\log r} = -\log_r(x)$$

ja väite on siten osoitettu.

Tehtävä 5

Määritellään kahden samassa perusjoukossa Ω määritellyn todennäköisyysfunktion P ja Q *Kullback–Leibler divergenssi* seuraavasti:

$$D_{KL}(P\|Q) := \sum_{x\in\Omega} P(x)\log_2\left(\frac{P(x)}{Q(x)}\right).$$

(Edellisissä demoissa näytit että $D_{KL}(P\|Q) \geq 0$.)
Osoita, että todennäköisyysjakauman $(\{a_1,\dots,a_k\}, F, P)$ entropia voidaan laskea kaavalla

$$H(P) = H(P_U) - D_{KL}(P\|P_U) = \log_2(k) - D_{KL}(P\|P_U),$$

missä P_U on tasainen todennäköisyysjakauma $P_U(a_j) = \frac{1}{k}$ joukossa Ω.

Ratkaisu:

Suoralla laskulla

$$\begin{aligned} H(P_U) - D_{KL}(P\|P_U) &= \log_2(k) - \sum_{a\in\Omega} P(a)\log_2\left(\frac{P(a)}{P_U(a)}\right) \\ &= \log_2(k) - \left(\sum_{a\in\Omega} P(a)\log_2(a) - \sum_{a\in\Omega} P(a)\log_2(P_U(a))\right) \\ &= \log_2(k) - \sum_{a\in\Omega} P(a)\log_2(a) + \log_2(1/k)\sum_{a\in\Omega} P(a). \end{aligned}$$

Kolmogorovin aksioomien nojalla

$$\sum_{a\in\Omega} P(a) = 1$$

ja entropian määritelmästä

$$H(P) = -\sum_{a\in\Omega} P(a)\log_2(a)$$

joten

$$\begin{aligned} H(P_U) - D_{KL}(P\|P_U) &= \log_2(k) - \sum_{a\in\Omega} P(a)\log_2(a) - \log_2(k) \\ &= -\sum_{a\in\Omega} P(a)\log_2(a) \\ &= H(P) \end{aligned}$$

ja väite on siten osoitettu.

Tehtävä 6

Anna esimerkki yhteisjakaumasta (X, Y) jossa

$$H(X|Y = b_k) > H(X)$$

jollakin b_k. (Tässä siis satunnaismuuttujiin X ja Y liittyvät alkeistapauksien joukot ovat $\{a_1, \dots, a_j\}$ sekä $\{b_1, \dots, b_i\}$.)
Osoita, että aina kuitenkin pätee

$$H(X|Y) \leq H(X).$$

Ratkaisu:

Määritellään yhteisjakauma (X, Y) siten, että

$$P(X, Y) = \begin{cases} 0, & X = a_1, Y = b_1 \\ 1/2, & X = a_2, Y = b_1 \\ 1/4, & X = a_1, Y = b_2 \\ 1/4, & X = a_2, Y = b_2 \end{cases}$$

Tällöin

$$H(X) = \frac{1}{4}\log_2(4) + \frac{3}{4}\log_2(4/3) \approx 0.81$$

ja

$$H(X|Y = b_2) = 1$$

eli $H(X|Y = b_2) = 1 > 0.82 > H(X)$. Osoitetaan vielä, että $H(X|Y) \leq H(X)$. Luentoprujun perusteella

$$I(X; Y) = D_{KL}(P(x, y) \| P(x)P(y))$$

ja

$$I(X; Y) = H(X) - H(X|Y).$$

Viime demoissa osoitettiin, että $D_{KL} \geq 0$, joten nämä yhdistämällä

$$I(X; Y) = H(X) - H(X|Y) \geq 0 \iff H(X|Y) \leq H(X).$$