

MATA2500 – Informaatioteoria. Kurssikoe 17.12.2019.

Jos aiot suorittaa tehtävän 2op laajuisena, vastaa **vain** perustehtäviin, jos aiot suorittaa tehtävän 4op laajuisina vastaa **sekä** perustehtäviin **että** syventäviin tehtäviin. Tentissä saa käyttää laskinta. Tentin liitteenä on lunttilappu jossa on sekalaisia kaavoja, sekä taulukko joistakin 2-kantaisen logaritmin arvoista.

Perustehtävät

Vastaa kahteen seuraavasta kolmesta kysymyksestä. Kukin kysymys on 6p arvoinen.

- (P1) Luennoitsija käy joka päivä lounaalla joko Maijassa, Piatossa tai Ilokivessä. Hänen todennäköisyytensä käydä Maijassa on $p_M = 0.5$, todennäköisyytensä käydä Piatossa on $p_P = 0.25$ ja todennäköisyytensä käydä Ilokivessä on $p_M = 0.25$.

Maijassa käydessään hän syö kasvislounaan todennäköisyydellä 0.75 ja lihalounaan todennäköisyydellä 0.25. Piatossa käydessään hän syö kasvislounaan todennäköisyydellä $7/8$ ja lihalounaan todennäköisyydellä $1/8$. Ilokivessä käydessään hän syö kasvislounaan todennäköisyydellä $1/16$ ja lihalounaan todennäköisyydellä $15/16$.

- (a) Muodosta todennäköisyysjakaumat lounaspaikan valinnalle X ja ruokavalion valinnalle Y sekä muodosta näiden yhteisjakauma (X, Y) .
 - (b) Laske entropia sekä yhteisjakaumalle (X, Y) että sen marginaalijakaumille X ja Y .
 - (c) Laske marginaalijakaumien X ja Y yhteisinformaatio $I(X; Y)$. (Yhden desimaalin tarkkuus riittää.)
- (P2) Tämä tehtävä liittyy Kolmogorov-kompleksisuuteen. Tehtävään on soveliaista antaa kuvaileva esseemäinen vastaus (n. puoli sivua) niin tahtoessaan, erityisesti jos ei ole käynyt kurssia *automaatit ja kielipit* (tai vastaavaa).
- (a) Kaverisi väittää kehittäneensä ohjelman, jolla voi kompressoida mitä tahansa tiedostoa kymmenellä prosentilla. Miten todistaisit taikka perustelisit hänelle, että tällaista ohjelmaa ei voi olla olemassa?
 - (b) Määrittele Kolmogorov-kompleksisuus.
- (P3) Tehtäväsi on suunnitella koodaus kohinattomalle binäärikanavalle, jonka lähetysnopeus on yksi merkki sekunnissa. Viestilähteenä sinulla on todennäköisyysjakauma $((A, B, C, D, E), (\frac{1}{16}, \frac{1}{16}, \frac{2}{16}, \frac{8}{16}, \frac{4}{16}))$.
- (a) Miten monta symbolia sekunnissa voit toivoa pystyväsi keskimäärin lähettämään tosi hyvällä koodauksella?
 - (b) Kuvaile jokin tapa, häviöllinen tai häviötön, päästä lähelle optimaalista rajaa.

Syventävät tehtävät

Vastaa kahteen seuraavasta kolmesta kysymyksestä. Kukin kysymys on 6p arvoinen.

- (S1) Olkoon (Y_0, Y_1) symmetrinen ε -kohinainen kanava, eli $Y_0 = ((0, 1), (1 - \varepsilon, \varepsilon))$, $Y_1 = ((0, 1), (\varepsilon, 1 - \varepsilon))$.
- (a) Laske kanavan Y kapasiteetti.
 - (b) Kuvaile jokin häviöllinen tai häviötön koodaus jolla voisit viestiä kanavalla keskimääräisellä virheellä, joka on alle $\varepsilon/2$.
- (S2) Olet mukana kisailuohjelmassa jossa edessäsi on kolme suljettua ovea. Yhden takana on palkinto, kahden muun takana ei mitään. Saat valita vapaasti oven, mutta ennenkuin avaat sen, kisan juontaja avaa kahdesta muusta ovesta sellaisen, jonka takana ei ole palkintoa. (Hän tietää palkinnon sijainnin.) Tämän jälkeen sinulle tarjotaan mahdollisuus vaihtaa ovea.
- Muodosta todennäköisyysjakauma joka kuvaa epävarmuuttasi tilanteessa, jossa olet valinnut alustavasti oven 1, mutta juontaja ei ole avannut vielä toista kahdesta jäljellejäävästä ovesta. Todennäköisyysjakaumassasi tulisi olla marginaalijakaumina palkinnon sijainti sekä juontajan avaaman oven numero. Laske näiden kahden satunnaismuuttujan yhteisinformaatio.
- (S3) Osoita, että todennäköisyysjakauman $X = ((a_1, \dots, a_k), P)$ entropia voidaan laskea kaavalla

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

missä P_U on tasainen todennäköisyysjakauma $P_U(a_j) = \frac{1}{k}$ joukossa Ω , ja $D_{KL}(P||P_U)$ tasajakauman P_U ja alkuperäisen jakauman P Kullback-Leibler -divergenssi.