JULIJ

**DATUM:** 1. 7., 2. 7., 3. 7., 4. 7.

**ŠT. UR:** 4 x 8 = 32

**DELO:** Že odkar sem začel delati na inštitutu, nisem odkril neke matematične definicije za informacijsko entropijo. Po nekaj brskanja po spletu sem naletel na članek – Pal. N. R in Pal S. K.: Entropy: A New Definition and its Applications (1991). Članek govori o vpeljavi definicije entropije, kjer za primerjavo uporabljajo Shannonovo entropijo. Entropijo rabimo vpeljati, ker hočemo računati posplošeno Jensen-Renyi divergenco (izračun temelji na entropiji). Nekaj dni sem si vzel, da sem natančno prebral omenjeni članek. Entropijo bom omenil tudi v zapiskih.

**DATUM:** 6. 7., 7. 7., 8. 7., 9. 7., 10. 7., 11. 7.

**ŠT. UR:** 6 x 8 = 48

**DELO:** V tem tednu sem se ukvarjal z ekvikvantalnimi stolpci. To so stolpci, v katerih je fiksno število podatkov. S tem sem že delal, moral pa sem stvari nekoliko obnoviti in dobro formulirati, da sem jih lahko zapisal v datoteko »O razliki porazdelitev«. Po testiranjih vseh treh metod za predstavitev podatkov (ekvidistantni stolpci, ekvikvantalni stolpci, Gaussian kernel density estimation) sem prišel do zanimivega zaključka: ekvidistantni stolpci niso nikoli optimalna izbira (razen mogoče uniformna porazdelitev, pa še takrat je ekvikvantalna rešitev skoraj ista). Pri gladkih porazdelitvah z neomejenimi nosilci je najboljša Gaussian kernel density estimation, pri porazdelitvah s špicami pa ekvikvanalni stolpci. Zanimivo pa je bilo, da se vse značilne lastnosti porazdelitve pokažejo že, če vzamemo 10 ekvikvantalnih stolpcev. To je odličen rezultat, ker je to zelo malo stolpcev, kar omogoča zelo hitro računanje.

**DATUM:** 13. 7., 14. 7., 15. 7., 16. 7., 17. 7., 18. 7.

**ŠT. UR:** 6 x 8 = 48

**DELO:** Ker sem prišel do konca 4. poglavja v zvezku »O razliki porazdelitev«, sem si vzel nekaj časa, da sem šel čez zvezek od začetka. Z jasno razdelanimi pojmi iz poglavja 4 sem bil bolj strog sam do sebe, zato sem našel kar nekaj napak, ki sem jih mogel odpraviti. Odkril sem napako v dokazu izreka: »Renyi divergenca je divergenca«, ki jo bom moral popraviti. Poleg tega sem našel nekaj pravopisnih napak in nekaj napak pri indeksiranju. V omenjeni zvezek želim vključiti še poglavje o entropiji in Jensenovi divergenci, ki je nekakšna posplošena divergenca (primerjamo več porazdelitev hkrati, ne le dveh). V tem tednu me je klical tudi Đani. Zmenila sva se, da mi pošlje literaturo, ki opisuje predstavitev podatkov z družino valčkov (kot razvoj v neskončno vrsto).

**DATUM:** 20. 7., 21. 7., 22. 7., 23. 7., 24. 7., 25. 7.

**ŠT. UR:** 6 x 8 = 48

**DELO:** Pisal sem poglavje o entropiji in posplošeni Jensen-Renyi divergenci. Napisal sem definicijo Renyi entropije glede na gostoto verjetnosti, nato pa vpeljal Renyi entropijo glede na histogram. Naredil sem nekaj izračunov Renyi entropij za beta porazdelitve, ki sem jih vključil v članek. Nadaljeval sem s poglavjem o družini Jensenovih divergenc, bolj podrobno pa sem opisal Jensen-Renyi divergenco (glede na gostoto verjetnosti in glede na histogram). Navedel sem še primer uporabe Jensen-Renyi divergence. Na koncu sem napisal še zaključek. Zvezek: »Divergenca kot merilo za razliko med porazdelitvami« (prej »O razliki porazdelitev«) prilagam v priponki.

**SKUPAJ ŠTEVILO UR: 176**

Priponka: Dokument »Divergenca kot merilo za razliko med porazdelitvami« (.pdf in .tex)