АСТРОСТАТИСТИКА - домаћи задаци

Правила:

- Домаће задатке је пожељно послати до назначених рокова, али није обавезно;
- Одрађени задаци се каче на гитхаб, JЕДИНО путем jupyter notebook документа;
- Сваки домаћи задатак је потребно детаљно прокоментарисати (у виду извјештаја, везано за то шта је одрађено у задатку, која је логика у позадини и слично). Извјештај треба качити на гитхаб у ПДФ формату, за сваки домаћи појединачно;
- Поени за сваки домаћи су подијељени на сљедећи начин:

```
70% - код;
30% - извјештај.
```

ПРВИ домаћи из астростатистике

1. задатак [40%]

- Генерисати N (гдје је N велики број) случајних узорака из униформне расподјеле у интервалу [a,b] и сачувати их као x;
- Направити хистограм ових узорака;
- Израчунати природни логаритам (ln) низа x и сачувати га као y;
- Направити нови хистограм за y;
- Користећи једначину за трансформацију расподјела вјероватноће, израчунати теоријску функцију густине вјероватноће (PDF) за y и плотовати је преко хистограма за y;
- Тражене плотове приказати један поред другог.

2. задатак [60%]

Ради се проширен Sleepy Beauty проблем. У недељу навече, организатори експеримента успављују љепотицу. Умјесто обичног, фер новчића, користи се пристрасан новчић, са вјероватноћом p(H)=p за главу и p(T)=1-p за писмо. У понедељак ујутру, експериментатори бацају новчић:

- ако падне глава, буде љепотицу само у понедељак,
- ако падне писмо, буде је N пута у различите дане, гдје је N случајна промјенљива (из Поасонове расподјеле, са параметром λ).

Када се пробуди, љепотица не зна који је дан и треба да процијени вјероватноћу да је "пала глава". Проблем треба ријешити у пајтону. Користити Бајесову формулу. Поасонова расподјела је дата као:

$$p(n=k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Додатно: пошто број буђења N, у случају писма, прати Поасонову расподјелу, ваља напоменути да λ представља очекиван број буђења ако је новчић пао на писмо. Узети да је $\lambda=2,3,4,$ а што се тиче вјероватноће пада главе, ставити да је p(H)=0.65. Може се искористити np.random.poisson. Приказати све тражене случајеве за λ , графички, један поред другог.

ДРУГИ домаћи из астростатистике

1. задатак [100%]

Анализирати магнетно поље Сунца - дат је FITS фајл у репозиторијуму.

- Плотовати читаву мапу магнетног поља;
- Приказати хистограм магнетног поља;
- Израчунати средњу вриједност, стандардну девијацију, медијану, искошеност и зашиљеност (за расподјелу која прати тај хистограм);
- Упоредити хистограм са гаусијаном;
- Издвојити 10% пиксела са највећим апсолутним вриједностима магнетног поља;
- Израчунати статистику за овај подскуп (средња вриједност, медијана, стандардна девијација);
- Упоредити статистичке мјере издвојених пиксела са статистиком цјелокупне мапе;
- Упоредити хистограм овог подскупа са гаусијаном и провјерити да ли постоје значајна одступања;
- Креирати профил магнетног поља дуж централне хоризонталне и вертикалне линије.

ТРЕЋИ домаћи из астростатистике

1. задатак [50%]

- Користити магнетно поље Сунца из претходног домаћег (исти FITS фајл);
- Подијелити мапу на четири квадратна подрегиона и за сваки квадрант израчунати просјечну вриједност и стандардну девијацију интензитета магнетног поља;
- Изабрати горњи лијеви и доњи десни квадрант израчунати Пирсонов коефицијент корелације између вриједности магнетног поља у ова два квадранта. Интерпретирати добијену вриједност.

2. задатак [50%]

Размотрити непрекидну функцију дефинисану на интервалу [0, 1]:

$$f(x) = x + \sin(5\pi x) + 1.$$

- Израчунати константу нормализације C тако да је

$$p(x) = \frac{f(x)}{C}$$

валидна густина расподјеле (тј. $\int_0^1 p(x) \, dx = 1$);

- Користити униформну расподјелу на [0,1]. Одредити константу M тако да за све $x \in [0,1]$ важи:

$$M g(x) \ge p(x);$$

- Имплементирати метод rejection sampling за узорковање из p(x) и наћи однос прихваћених узорака;
- Израчунати кумулативну функцију F(x):

$$F(x) = \int_0^x p(t) \, dt;$$

- Нумерички доћи до функције $F^{-1}(u)$, гдје је $u \sim \mathrm{Uniform}(0,1);$
- Генерисати узорке примјеном F^{-1} на униформно расподијељене бројеве;
- За обје методе плотовати хистограме узорака и упоредити их са теоријском функцијом густине p(x).

ЧЕТВРТИ домаћи из астростатистике

1. задатак [100%]

- Користити фајл *bhm.npy* који садржи масе црних рупа (у масама Сунца);
- Израчунати стандардну девијацију, медијану и коефицијент асиметрије;
- Процијенити неодређености у ове три статистике, користећи:
 - а) Bootstrap методу (са 10000 реузорковања),
 - б) Jackknife методу (избацивање по два податка, leave two out);
- Зашто је за извршавање *Jackknife* методе потребно значајно више времена?
- Упоредити резултате визуелно и закључити која метода даје поузданију процјену, за сваку статистику посебно;
- Убацити десет екстремних вриједности ($M_{\rm BH}=1000\,M_\odot$) у податке, и закључити како то утиче на све горе наведено.

Црне рупе се могу формирати на различите начине (канали формирања) и класификују се у популације. Претпостављамо да свака популација има гаусовску расподјелу, а да је укупна расподјела мјешавина више гаусовских:

$$p(m) = \sum_{j=1}^{N} \alpha_j \mathcal{N}(m|\mu_j, \sigma_j), \quad \sum_{j=1}^{N} \alpha_j = 1.$$

- На основу плота расподјеле маса црних рупа, наћи највише пикове;
- Користити sklearn.mixture.GaussianMixture за фитовање модела за различит број компоненти (N=1,2,...,10);
- За сваку вриједност N утврдити тежински фактор α_j , просјечне масе μ_j и стандардне девијације σ_i ;
- Плотовати хистограме појединачних гаусовских компоненти које представљају добијене расподјеле;
- Испитати стабилност добијених параметара у зависности од броја компоненти, те објаснити шта то указује на могуће канале формирања црних рупа.

ПЕТИ домаћи из астростатистике

1. задатак [100%]

Црна рупа масе \mathcal{M} и спина¹ $\chi \in [0,1]$, има несводљиву (енг. *irreducible*) масу:

$$\mathcal{M}_{\mathrm{irr}} = \mathcal{M} \sqrt{rac{1+\sqrt{1-\chi^2}}{2}}.$$

Користимо $f = \mathcal{M}_{irr}/\mathcal{M}$. Имамо поновљена мјерења од \mathcal{M} и χ , претпостављамо да је спин униформно расподијељен у горенаведеном интервалу, а да маса прати нормалну расподјелу $\mathcal{N}(\mu, \sigma)$.

- Користити вриједности $\sigma = 0.02$ и $\mu = 1$ (све масе у проблему су изражене у јединицама μ);
- Плотовати добијену PDF функцију за $M_{\rm irr}$. Користити Скотово или Фридман-Дијаконисово правило.
- Плотовати расподјелу $M_{\rm irr}$ помоћу $Kernel\ Density\ Estimation\ (KDE)$ методе.
- Израчунати Колмогоров-Смирнов дистанцу D између $M_{\rm irr}$ и f као функцију од σ . Коментарисати границе за велике и мале вриједности σ .
- Израчунати Колмогоров-Смирнов дистанцу D између $M_{\rm irr}$ и M као фунцију од σ . Коментарисати границе за велике и мале вриједности σ .
- Израчунати PDF од f. Може се користити формула:

$$\pi(f) = 2 \frac{2f^2 - 1}{\sqrt{1 - f^2}}.$$

Провјерити да ли је $\pi(f)$ нормализовано.

- Израчунати PDF од $M_{\rm irr}$. Корисна је формула:

$$p(M_{\rm irr}) = \frac{\sqrt{2/\pi}}{\sigma} \int_{1/\sqrt{2}}^{1} e^{-\frac{\left(M_{\rm irr}/f - \mu\right)^2}{2\sigma^2}} \frac{2f^2 - 1}{f\sqrt{1 - f^2}} df.$$

 $^{^{1}}$ Важи формула $\chi=\frac{cJ}{GM^{2}},$ гдје је Jуга
они момент.

ШЕСТИ домаћи из астростатистике

1. задатак [100%]

Треба написати кратак извјештај о проблему који је описан на овом линку. Нема потребе репродуковати проблем у *пајтону*.