

Entropia Rényi'ego - zadania

Piotr Migdał pmigdal@gmail.com

7 czerwca 2015 (poprawione: 13 czerwca 2015)

Zadania głównie z entropii Shannona. Jeśli czegoś nie wiesz lub jest niejasne - nie bój się pytać, chętnie pomogę!

Przeczytanie dwóch pierwszych rozdziałów z *J. Thomas and T. Cover, Elements of Information Theory* (linki na stronie warsztatów) nie wymagane, ale mile widziane. :)

Najważniejszym (\neq najtrudniejszym) zadaniem jest 5.

Zadanie 1 - jaka jest entropia Shannona

(w bitach, jako liczbę)

- 1.1. Rzutu kostką do gry?
- 1.2. Kolejności ułożenia 52 kart w talii?
- 1.3. Liczby wykonanych rzutów uczciwą monetą, gdy rzucamy nią tyle razy aż wypadnie orzeł?

Zadanie 2 - własności entropii Shannona

- 2.1. Pokaż, że w przypadku rozkładu prawdopodobieństwa $(p, 1-p)$ entropia Shannona osiąga maksimum gdy $p = 1/2$. (W wersji z \star - a jak to wygląda dla rozkładu o n możliwościach?)
- 2.2. Pokaż, że entropia zmiennych niezależnych to suma entropii tych zmiennych. (W wersji z \star - pokaż, że tak się dzieje wtedy i tylko wtedy.)

Zadanie 3 - entropia Shannona a niepewność

W rozkładzie (p_1, p_2, \dots, p_n) niech p_i oznacza prawdopodobieństwo, że cząstka jest na i -tej pozycji.

- 3.1. Dla $n = 2$ wyraż zależność entropii i wariancji od p_1 . (Analitycznie + załącz wykres.)
- 3.2. (\star) Dla $n = 5$ pozycji przeprowadź symulację losującą różne rozkłady prawdopodobieństwa i przedstaw wykres punktowy wariancji vs entropii Shannona. (Załącz też kod.)

Zrób to w dowolnym języku. Jeśli nie wiesz od czego zacząć, polecam środowisko [IPython Notebook](#) do Python 3 instalowanego z pakietu [Anaconda](#). Innym dobrym wyborem może być [Julia](#).

Zadanie 4 - entropia różniczkowa

Człowiek czasem spotyka taki napis:

$$-\int p(x) \ln(p(x)) dx$$

{#eq:description}

który jest zwany entropią różniczkową (Shannona). Jest ściśle związany ze “zwykłą” entropią, ale ma niektóre inne własności - np. może być ujemne.

Pokaż (na bezpośrednich przykładach), że ta wielkość może być:

- 4.1. ujemna,
- 4.2. zerowa,
- 4.3. dodatnia,
- 4.4. (*) $+\infty$,
- 4.5. (*) $-\infty$.

Zadanie 5 - interpretacje i zastosowania

Wyniki wyborów prezydenckich z pierwszej tury są tu: http://prezydent2015.pkw.gov.pl/319_Pierwsze_glosowanie. Traktując to jak rozkład prawdopodobieństwa zagłosowania na danego kandydata:

- 5.1. Wykreśl entropię Rényi’ego dla stopnia α od 0 do 20 (plus zaznacz asymptotę w ∞).
- 5.2. Jak się zmieni wykres, jak:

- obetniemy do trzech najbardziej popularnych kandydatów?
- przedstawimy prawdziwe wyniki tj. uwzględniające wstrzymanie się od głosu?

(W obu przypadkach nie zapomnij o normalizacji prawdopodobieństwa! Wykresy z 5.2. najlepiej przedstawić w postaci nałożonej na 5.1.)

5.3. Zinterpretuj H_0 , H_2 oraz H_∞ . Możesz (a nawet musisz :)) przekształcić owe H_α przy pomocy jakichś funkcji (tj. rzeczy w stylu $\sin(H_2^4 + 1)$). Czym prostsza interpretacja tym lepiej.

Jak nie wiesz w czym rysować, to dobrymi wyborami są wspomiane Python (np. [pandas](#)) i Julia ([Gadfly](#)), ale niesamowity jest [ggplot2](#) w R - [Quick introduction to ggplot2](#).