Beberapa Distribusi Diskrit

Ali Akbar Septiandri

Universitas Al-Azhar Indonesia

aliakbars@live.com

March 26, 2019

Selayang Pandang

Ulasan

Minggu lalu...

- Peubah acak, ruang sampel, kejadian
  Jika dilempar sebuah dadu, $p(X = 5) = \ldots$

Minggu lalu...

- Peubah acak, ruang sampel, kejadian
  Jika dilempar sebuah dadu, $p(X = 5) = ...$
- Ekspektasi dan variansi
  $\mathbb{E}[X] = ...$ dan $Var[X] = ...$

Minggu lalu...

- Peubah acak, ruang sampel, kejadian
  Jika dilempar sebuah dadu, $p(X = 5) = ...$
- Ekspektasi dan variansi
  $\mathbb{E}[X] = ...$ dan $Var[X] = ...$
- Marginal probability, conditional probability
  sum rule, product rule, chain rule

Minggu lalu...

- Peubah acak, ruang sampel, kejadian
  Jika dilempar sebuah dadu, $p(X = 5) = ...$
- Ekspektasi dan variansi
  $\mathbb{E}[X] = ...$ dan $Var[X] = ...$
- Marginal probability, conditional probability
  sum rule, product rule, chain rule
- Bayes' rule
  $p(y|x) = \frac{p(x|y)p(y)}{p(x)}$

Ekspektasi dan Variansi

Ekspektasi

$$\mathbb{E}[X] = \sum_{x} f(x)p(x)$$

Ekspektasi dan Variansi

Ekspektasi

$$\mathbb{E}[X] = \sum_{x} f(x)p(x)$$

Variansi

$$Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Tunjukkan!

Sudah coba soal latihan?

Sudah enroll ke e-learning?

Distribusi Uniform

Lemparan Dadu

Peubah Acak Seragam

- Semua bernilai sama peluangnya dalam interval tertentu
- Dituliskan sebagai $X \sim Unif(a, b)$
- Bisa berupa diskrit maupun kontinu
- PMF:

$$p(x) = \begin{cases} \frac{1}{b-a+1} & , if \;\; x \in [a, b] \\ 0 & , otherwise \end{cases}$$

Ekspektasi dan Variansi

Ekspektasi

$$\mathbb{E}[X] = \frac{a+b}{2}$$

Variansi

$$Var[X] = \frac{(b-a+1)^2 - 1}{12}$$

German tank problem

Distribusi Bernoulli

Definisi

- X adalah RV yang bisa bernilai 0 atau 1

Definisi

- X adalah RV yang bisa bernilai 0 atau 1
- Nilai 1 dapat diartikan sebagai "sukses", e.g. muncul "angka" dari koin, muncul muka 4 atau 5 dari dadu, ads akan diklik

Definisi

- X adalah RV yang bisa bernilai 0 atau 1
- Nilai 1 dapat diartikan sebagai "sukses", e.g. muncul "angka" dari koin, muncul muka 4 atau 5 dari dadu, ads akan diklik
- Jika $p(X = 1|\theta) = \theta$ dan mengakibatkan $p(X = 0|\theta) = 1 - \theta$, maka

Definisi

- X adalah RV yang bisa bernilai 0 atau 1
- Nilai 1 dapat diartikan sebagai "sukses", e.g. muncul "angka" dari koin, muncul muka 4 atau 5 dari dadu, ads akan diklik
- Jika $p(X = 1|\theta) = \theta$ dan mengakibatkan $p(X = 0|\theta) = 1 - \theta$, maka
- X mengikuti distribusi Bernoulli

Distribusi Bernoulli

PMF

$$X \sim Bernoulli(\theta)$$

$$p(x) = \begin{cases} \theta & , x = 1 \\ 1 - \theta & , x = 0 \end{cases}$$

dapat diringkas sebagai

$$p(x) = \theta^x (1 - \theta)^{1-x}$$

Model Bernoulli

Example

Data: 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 1 1 0 1.

Tiga hipotesis:

- $\mathcal{M} = 1$ - dari koin setimbang $1 = H$, $0 = T$
- $\mathcal{M} = 2$ - dari lemparan dadu $1 = 1$, $0 = 2, 3, 4, 5, 6$
- $\mathcal{M} = 3$ - dari koin yang keduanya muka $1 = H$, $0 = T$

Likelihood

- $p(\mathcal{D}|\mathcal{M})$, yaitu probabilitas melihat data $\mathcal{D}$ jika diberikan distribusi (atau model) $\mathcal{M}$
- Merupakan hasil perkalian dari peluang yang menghasilkan tiap titik dalam data

$$L(\mathcal{M}) = p(\mathcal{D}|\mathcal{M}) = \prod_{i=1}^{N} p(x_i|\mathcal{M})$$

- Kita menggunakan asumsi *independent and identically distributed* (i.i.d.)

Model Bernoulli

Example

Data: 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 1 1 0 1.

Likelihood of data. Jika N_1 = jumlah 1, N_0 = jumlah 0, dengan $N = N_0 + N_1$:

$$\prod_{i=1}^{N} p(x_i|\mathcal{M}) = p(1|\mathcal{M})^{N_1} p(0|\mathcal{M})^{N_0}$$

maka

- $\mathcal{M} = 1 : L(\mathcal{M}) = 0.5^{20} = 9.5 \times 10^{-7}$
- $\mathcal{M} = 2 : L(\mathcal{M}) = (\frac{1}{6})^9 (\frac{5}{6})^{11} = 1.3 \times 10^{-8}$
- $\mathcal{M} = 3 : L(\mathcal{M}) = 1^9 0^{11} = 0$

Maximum Likelihood Estimation

- Coba berbagai model $\mathcal{M}$ yang dapat memaksimalkan nilai *likelihood*, i.e. *maximum likelihood estimation*
- Dalam kasus distribusi Bernoulli

$$L(\mathcal{M}) = p(X|\theta) = \prod_{i=1}^{N} \theta^{x_i} (1-\theta)^{1-x_i}$$

- Bagaimana?

Beberapa trik...

1. Terapkan fungsi logaritma
2. Cari turunan parsial pertama
3. Atur sama dengan nol

Coba cari MLE-nya!

Bernoulli MLE

$$\hat{\theta} = \frac{N_1}{N_0 + N_1}$$

Distribusi Binomial

Distribusi Binomial

- Didapatkan dari n percobaan Bernoulli independen
- Distribusinya menunjukkan jumlah kemunculan 1
- X dapat bernilai $0, 1, 2, ..., n$
- Jika
$$p(X = r|\theta) = \binom{n}{r}\theta^r(1-\theta)^{n-r},$$
maka X mengikuti distribusi Binomial

Catatan tentang Binomial

- $X \sim Binom(n, \theta)$
- $\mathbb{E}[X] = n\ \theta$
- $Var(X) = n\ \theta(1 - \theta)$
- $Ber(\theta) = Bin(1, \theta)$

Contoh

1. Anda melempar koin setimbang sebanyak 20 kali. Berapa peluang muncul gambar 10 kali?

Contoh

1. Anda melempar koin setimbang sebanyak 20 kali. Berapa peluang muncul gambar 10 kali?
2. Anda melakukan 20 lemparan koin lagi, tetapi kali ini dengan koin yang bias dengan $\theta = 0.2$ dan kemunculan gambar sebagai "sukses". Berapa peluangnya muncul gambar 10 kali?

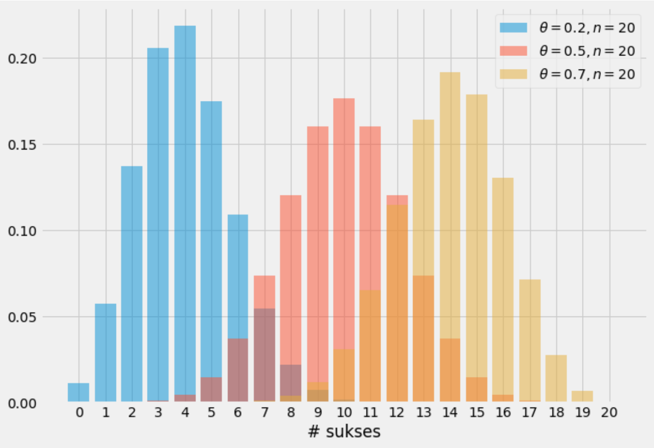
Contoh

1. Anda melempar koin setimbang sebanyak 20 kali. Berapa peluang muncul gambar 10 kali?
2. Anda melakukan 20 lemparan koin lagi, tetapi kali ini dengan koin yang bias dengan $\theta = 0.2$ dan kemunculan gambar sebagai "sukses". Berapa peluangnya muncul gambar 10 kali?
3. Berapa peluang minimal ada satu kemunculan gambar?

Contoh

1. Anda melempar koin setimbang sebanyak 20 kali. Berapa peluang muncul gambar 10 kali?
2. Anda melakukan 20 lemparan koin lagi, tetapi kali ini dengan koin yang bias dengan $\theta = 0.2$ dan kemunculan gambar sebagai "sukses". Berapa peluangnya muncul gambar 10 kali?
3. Berapa peluang minimal ada satu kemunculan gambar?
4. Dengan kasus seperti sebelumnya. Berapa kali kemunculan gambar yang Anda harapkan?

Grafik PMF



Gambar: Probability mass function (PMF) dari distribusi Binomial

Ikhtisar

- Distribusi uniform
- Distribusi Bernoulli
- Distribusi Binomial, dapat dianggap sebagai *Bernoulli process*
- Maximum Likelihood Estimation

Topik Tambahan (non-examinable)

- Distribusi Poisson
- Distribusi Geometrik
- Distribusi Negative Binomial

Pertemuan Berikutnya

- Distribusi Beta
- Distribusi Gaussian
- Multivariate Gaussian

Office hours minggu depan di
hari Senin, 08.00-09.00

Referensi

Will Monroe (Jul. 2017)
Bernoulli and Binomial Random Variables
http://web.stanford.edu/class/archive/cs/cs109/cs109.1178/lectureHandouts/070-bernoulli-binomial.pdf

Chris Williams (Sep. 2015)
Probability - Machine Learning and Pattern Recognition
https://www.inf.ed.ac.uk/teaching/courses/mlpr/2015/

Terima kasih