Random Variables

W. M. Song 桑慧敏
Tsing Hua Univ. 清華大學

2015.09.23

Practice Problems

1. List two well-known random variables (有名字的隨機變數):
   ____________ and ____________.
2. If we know ________, we can compute the probability of any event relating to the random variable X.
3. 解釋歌詞意思

樣本點, 實數值, 隨機變數結連理

cdf, pdf, 身份證件辨唯一

期望值, 變異數, 又名一二階動差

馬可夫, 柴比雪夫, 神龍現尾真稀奇

Random Variable (Song)

隨機變數基本概念 (三輪車調)

- 樣本點, 實數值, 隨機變數結連理
- cdf, pdf, 身份證件辨唯一
- 期望值, 變異數, 又名一二階動差
- 馬可夫, 柴比雪夫, 神龍現尾真稀奇

- sample point, real number, random variable connects them
- cdf. pdf, ID the role they play
- E(X), V(X), 1st and 2nd Moments
- Markov, Chebyshev, P of tail events so easy.

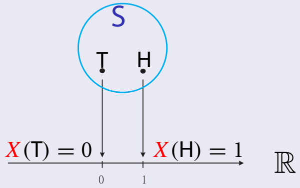
- 作業是熟讀本章的內容, 然後修改歌詞。以能闡釋該章核心精神為評分標準。

樣本點, 實數值, 隨機變數結連理

Ex. Flip a fair coin

- Outcomes "H" or "T"
- Compute P("H")
- Sample space: S= { H, T }

X: $s \to r, s \in$ S, $r \in \mathbb{R}$



- Q: List one Advantage of Using the R.V. X
- A: Quantitative (量化)

Express “Events” via R.V.

Type	Event, E (完整)	Event (縮寫)	P(E) (完整)	P(E) (縮寫)
1	$\{s : X(s) \leq x\}$	$\{X \leq x\}$	$\mathsf{P}(\{X \leq x\})$	$\mathsf{P}(X \leq x)$
2	$\{s : X(s) < x\}$	$\{X < x\}$	$\mathsf{P}(\{X < x\})$	$\mathsf{P}(X < x)$
3	$\{s : X(s) \geq x\}$	$\{X \geq x\}$	$\mathsf{P}(\{X \geq x\})$	$\mathsf{P}(X \geq x)$
4	$\{s : X(s) = x\}$	$\{X = x\}$	$\mathsf{P}(\{X = x\})$	$\mathsf{P}(X = x)$
5	$\{s : X(s) > x\}$	$\{X > x\}$	$\mathsf{P}(\{X > x\})$	$\mathsf{P}(X > x)$
6	$\{s : a \leq X(s) \leq b\}$	$\{a \leq X \leq b\}$	$\mathsf{P}(\{a \leq X \leq b\})$	$\mathsf{P}(a \leq X \leq b)$
7	$\{s : a \leq X(s) < b\}$	$\{a \leq X < b\}$	$\mathsf{P}(\{a \leq X < b\})$	$\mathsf{P}(a \leq X < b)$
8	$\{s : a < X(s) \leq b\}$	$\{a < X \leq b\}$	$\mathsf{P}(\{a < X \leq b\})$	$\mathsf{P}(a < X \leq b)$
9	$\{s : a < X(s) < b\}$	$\{a < X < b\}$	$\mathsf{P}(\{a < X < b\})$	$\mathsf{P}(a < X < b)$

- Q: List more advantages of Using a Random Variable
- A: (1) Quantitative (量化)
- A: (2) Simplify the expression for events (ex. Flip a fair coin twice, event that one “H” appears: {(T,H), (H,T)}. Can we simply this expression?
- A: (3) All events can be expressed as one of the above 9 types or their combinations

隨機變數 X, 證件, 屬性

Identification (ID, 證件)

- cdf: $F_X(x) = \mathrm{P}(X \leq x)$
- pmf or pdf:

$$f_X(x) = \begin{cases} F_X(x) - F_X(x^-) = \mathrm{P}(X = x), & \text{若 } X \text{ 為離散型隨機變數} \\ \dfrac{dF_X(x)}{dx}, & \text{若 } X \text{ 為連續型隨機變數} \end{cases}$$

- mgf: $\mathrm{M}_X(t) = \mathrm{E}(e^{tx})$

Characteristics (屬性)

- Moments (動差): Expected value, variance, skewness, and kurtosis,...
- quartiles (四分位數)
- Percentiles (百分位數)
- Tail events probability (尾端機率) : Markov and Chebyshev's Inequalities

- 將有關隨機變數的課題分成兩類: 證件與屬性, 你的看法呢?

ID 1: $F_X(x)$, cdf of X

Definiton of $F_X(x)$

$$F_X(x) = \mathrm{P}(X \leq x)$$

type	P(E)	Via $F_X(x)$
1	$\mathrm{P}(X \leq x)$	$F_X(x)$
2	$\mathrm{P}(X < x)$	$F_X(x^-)$
3	$\mathrm{P}(X \geq x) = 1 - \mathrm{P}(X < x)$	$1 - F_X(x^-)$
4	$\mathrm{P}(X > x) = 1 - \mathrm{P}(X \leq x)$	$1 - F_X(x)$
5	$\mathrm{P}(X = x) = \mathrm{P}(X \leq x) - \mathrm{P}(X < x)$	$F_X(x) - F_X(x^-)$
6	$\mathrm{P}(a \leq X \leq b) = \mathrm{P}(X \leq b) - \mathrm{P}(X < a)$	$F_X(b) - F_X(a^-)$
7	$\mathrm{P}(a < X \leq b) = \mathrm{P}(X \leq b) - \mathrm{P}(X \leq a)$	$F_X(b) - F_X(a)$
8	$\mathrm{P}(a \leq X < b) = \mathrm{P}(X < b) - \mathrm{P}(X < a)$	$F_X(b^-) - F_X(a^-)$
9	$\mathrm{P}(a < X < b) = \mathrm{P}(X < b) - \mathrm{P}(X \leq a)$	$F_X(b^-) - F_X(a)$

- Q: Why can $F_X(x)$ play the role of "ID"?
- A: With $F_X(x)$, we can compute the probability of any event relating to X

ID 2: $f_X(x)$, pmf or pdf of X

Definiton of $f_X(x)$

$$f_X(x) = \begin{cases} F_X(x) - F_X(x^-) = \mathrm{P}(X = x), & \text{若 } X \text{ 為離散型隨機變數} \\ \dfrac{dF_X(x)}{dx}, & \text{若 } X \text{ 為連續型隨機變數} \end{cases}$$

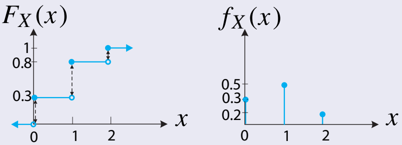
type	P(E)	discrete	continuous
1	$\mathrm{P}(X \le x)$	$\sum_{y \le x} f_X(y)$	$\int_{-\infty}^{x} f_X(y)dy$
2	$\mathrm{P}(X < x)$	$\sum_{y < x} f_X(y)$	$\int_{-\infty}^{x} f_X(y)dy$
3	$\mathrm{P}(X \ge x)$	$1 - \sum_{y < x} f_X(y)$	$1 - \int_{-\infty}^{x} f_X(y)dy$
4	$\mathrm{P}(X > x)$	$1 - \sum_{y \le x} f_X(y)$	$1 - \int_{-\infty}^{x} f_X(y)dy$
5	$\mathrm{P}(X = x)$	$f_X(x)$	0 (why?)
6	$\mathrm{P}(a \le X \le b)$	$\sum_{a \le y \le b} f_X(y)$	$\int_a^b f_X(y)dy$
7	$\mathrm{P}(a < X \le b)$	$\sum_{a < y \le b} f_X(y)$	$\int_a^b f_X(y)dy$
8	$\mathrm{P}(a \le X < b)$	$\sum_{a \le y < b} f_X(y)$	$\int_a^b f_X(y)dy$
9	$\mathrm{P}(a < X < b)$	$\sum_{a < y < b} f_X(y)$	$\int_a^b f_X(y)dy$

- A: With $f_X(x)$, we can compute the probability of any event

$F_X(x)$ and $f_X(x)$

discrete R.V.s

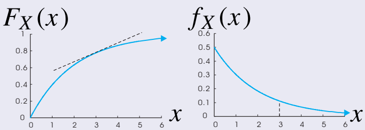
- $f_X(x) = F_X(x) - F_X(x^-) = \mathrm{P}(X = x)$
- $F_X(a) = \sum_{x \le a} f_X(x)$



(a) cdf (b) pmf

continous R.V.s

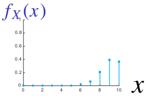
- $f_X(x) = \dfrac{dF_X(x)}{dx}$
- $F_X(a) = \int_{x \le a} f_X(x)dx$



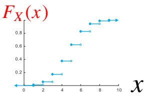
(a) cdf (b) pdf

- Discussion. The use of plots

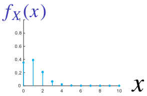
Match $F_X(x)$ with $f_X(x)$



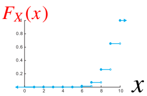
(a)



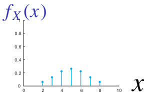
(1)



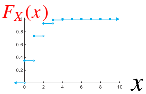
(b)



(2)



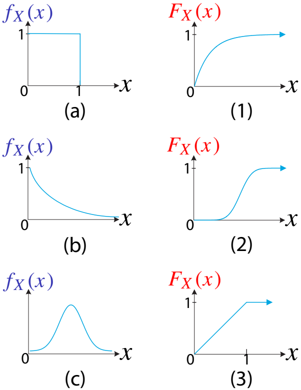
(c)



(3)

- Match: $(a, 2), (b, 3), (c, 1)$

Match $F_X(x)$ with $f_X(x)$ (conti.)



- Match: $(a, 3), (b, 1), (c, 2)$

期望值, 變異數, 又名一二階動差

Characteristics (屬性) of random variables

- Moments (動差): $E(X^k)$ or $E[(X-\mu_X)^k]$; Expected value, variance, skewness, and kurtosis,...
- quartiles (四分位數): $q_i = \mathrm{P}(X \leq q_i) \geq 0.25i, i = 1, 2, 3$
- Percentiles (百分位數)
- Tail events probability (尾端事件機率) : Markov and Chebyshev's Inequalities
- Process capacity measures (製程能力指標): $C_p, C_{pk}, \ldots$ (text 8.3)

馬可夫，柴比雪夫，神龍現尾真稀奇

- Markov's Inequality: Suppose that Y is a non-negative random variable and $\mathrm{E}(Y) < \infty$. For any positive real number $a > 0$, we have

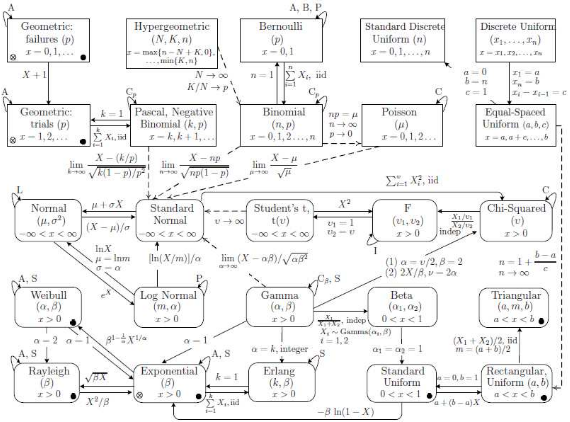
$$\mathrm{P}(Y \geq a) \leq \frac{\mathrm{E}(Y)}{a}$$

- Chebyshev's Inequality: Let X be any random variable with $\mathrm{E}(X) < \infty$, $\sigma_X^2 < \infty$. For any $k > 0$,

$$\mathrm{P}(|X - \mu_X| \geq k\sigma_X) \leq 1/k^2$$

- With Markov's and Chebyshev's Inequality, but without cdf or pdf, we can get the probability upper bounds of some events.

Well-Known Random Variables



- Reference: Relationships among 25 Distributions (Song, W.-M. T. and Chen, Yi-chun (2011), Eighty Univariate Distributions and their Relationships Displayed in a Matrix Format, IEEE Transactions on Automatic Control, Vol. 56, No. 8, pp. 1979-1984.)

More questions

- 1. 桑老師試作的 "隨機變數 (三輪車調)", 大家試著唱唱看, 並想一想此歌詞是否有幫助你更了解本章涵意?
- 2. 舉例說明 "隨機變數" 的功用?
- 3. 累積分配函數 $F_X(x)$ 與分配函數 $f_X(x)$ 之關係為何? 分別以公式與圖形說明之。
- 4. 符號 $F_X(x)$ 與 $f_X(x)$ 中下標示大寫, 括號內是小寫, 為何?
- 5. 累積分配函數與分配函數既然有一對一的關係, 何時特別需要使用分配函數? 何時特別需要使用累積分配函數? 舉例說明之。
- 6. 列出一個有名字的隨機變數。(hint: 如果隨機變數 X 服從常態分配, 則稱隨機變數 X 為常態隨機變數)
- 7. If we know cdf F_X, or pdf f_X, we can compute the probability of any event relating to the random variable X.
- 參考書: 桑慧敏, 機率與統計原理第三章, McGrawHill 出版, 2007