



King Mongkut's University of Technology Thonburi  
 Faculty of Engineering, Department of Computer Engineering  
 CPE 213 Data Models, 2/2020

LAB Lecture 7: Modeling statistical distribution  
 Assign Date: 26 Mar 2021 Due Date: 1 April 2021

### 1) Select one numeric column of your data

Dataset ที่ทำการเลือกมาคือ Dataset เกี่ยวกับ Heart Attack Analysis & Prediction Dataset โดยจะเป็น Dataset ที่เกี่ยวข้องกับหัวใจวาย และได้ทำการเลือก Column เป็น thalach คือ Column ที่อธิบาย Heart rate สูงสุดของตัวอย่างนั้น

A data.frame: 303 x 14

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<dbl>	<int>	<int>	<int>	<int>
63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
57	1	2	150	168	0	1	174	0	1.6	2	0	2	1
54	1	0	140	239	0	1	160	0	1.2	2	0	2	1
48	0	2	130	275	0	1	139	0	0.2	2	0	2	1

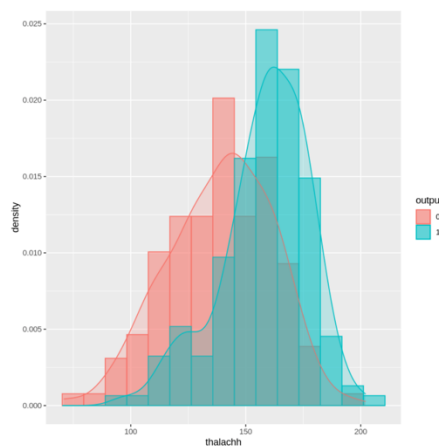


King Mongkut's University of Technology Thonburi  
 Faculty of Engineering, Department of Computer Engineering  
 CPE 213 Data Models, 2/2020

LAB Lecture 7: Modeling statistical distribution

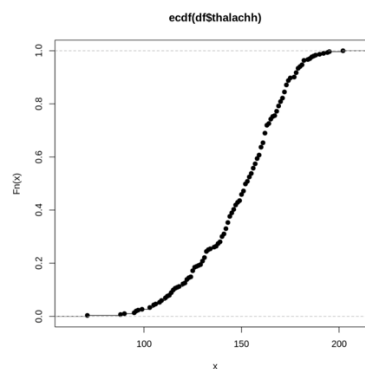
Assign Date: 26 Mar 2021 Due Date: 1 April 2021

## 2) Plot probability distribution function



จากกราฟ PDF เราจะเห็นได้ว่าการ การกระจายตัวนั้นจะเป็นแบบ Binomial ก็คือ โอกาสของคนที่เป็นโรคหัวใจวายจะมี Heart rate สูงสุดอยู่ที่เท่าใดได้บ้าง (output = 1)

## 3) Plot CDF



จากกราฟ CDF นี้จะแสดงค่าความน่าจะเป็น ของ Heart rate สูงสุด ในแต่ละค่าที่จะเกิดได้ ตัวอย่างเช่น สุ่มความน่าจะเป็นที่เกิดขึ้นได้ 0.4 เราจะได้ค่าที่ควรออกมาคือ 150 โดยประมาณ



King Mongkut's University of Technology Thonburi  
Faculty of Engineering, Department of Computer Engineering  
**CPE 213 Data Models, 2/2020**

LAB Lecture 7: Modeling statistical distribution  
Assign Date: 26 Mar 2021 Due Date: 1 April 2021

---

#### 4) Sampling 10 values from your distribution

89.02338%:	174.850596223492	17.6673%:	126 8.921542%:	114.943056755699	36.29466%:	144 44.14559%:	150 46.56627%:	151 48.46495%:
152 29.73941%:	140 34.97247%:	143 27.95363%:	139.419958413113					

นี่คือ Sampling ที่ถูกสุ่มมาจาก Data ของ thalach โดยจะบอก ค่าความน่าจะเป็นและ ค่าที่ควรจะได้