Statistik

für Betriebswirtschaft und internationales Management

Sommersemester 2015

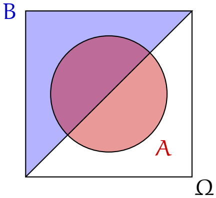
Prof. Dr. Stefan Etschberger
Hochschule Augsburg

Bedingte Wahrscheinlichkeiten

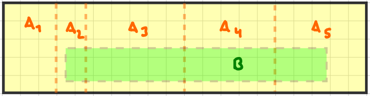
- Wahrscheinlichkeit von A hängt von anderem Ereignis B ab. (B kann zeitlich vor A liegen, muss aber nicht!)
- **Beispiel:** Wahrscheinlichkeit für Statistiknote hängt von Mathenote ab.
- Formal:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

- Im Venndiagramm:



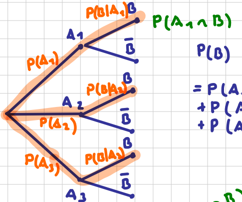
Es gelte: $\bigcup_{i=1}^{n} A_i = \Omega$ und $A_i \cap A_j = \{\}$ (für $i \neq j$)



$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum P(B|A_i) \cdot P(A_i)$$

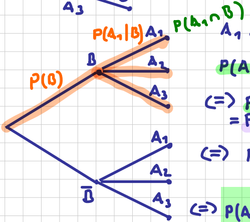
Satz von der totalen Wahrscheinlichkeit

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$



$P(B)$

$= P(A_1) \cdot P(B|A_1)$
$+ P(A_2) \cdot P(B|A_2)$
$+ P(A_3) \cdot P(B|A_3)$



$A_1 \cap B$

$P(A_1 \cap B) = P(B \cap A_1)$

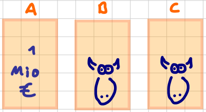
$\Leftrightarrow P(A_1|B) \cdot P(B)$
$= P(B|A_1) \cdot P(A_1)$

$\Leftrightarrow P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)}$

$$\Leftrightarrow P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{n} P(A_j) \cdot P(B|A_j)}$$

Satz von Bayes

Beispiel: Ziegenproblem



Quizshow: Kandidat muss sich für eine Tür entscheiden (hinter 1 Tür: 1 Mio)

Nach Entscheidung öffnet der Moderator eine „Nieten"-Tür; danach: Kandidat darf wechseln

Soll er wechseln?

0	A	Wechseln ist schlecht
1	B	Wechseln ist besser
6	C	Egal
6	D	Ich kenne das

Ereignisse: A, B, C: Die Mio ist hinter Tür A, B, C

M_A, M_B, M_C: Moderator öffnet die Tür A, B, C

Kandidat wählt A:

$P(M_B|A) = \frac{1}{2}$
$P(M_B|B) = 0$
$P(M_B|C) = 1$

$P(M_B) = P(M_B|A) \cdot P(A)$
$+ P(M_B|B) \cdot P(B)$
$+ P(M_B|C) \cdot P(C)$

$\left.\right\} P(A) = P(B) = P(C) = \frac{1}{3}$

$= \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2}$

Kandidat bleibt bei Entscheidung:

$$P(A \mid M_B) = \frac{P(M_B \mid A) \cdot P(A)}{P(M_B)} = \frac{1/2 \cdot 1/3}{1/2} = 1/3$$

Kandidat wechselt Tür:

$$P(C \mid M_B) = \frac{P(M_B \mid C) \cdot P(C)}{P(M_B)} = \frac{1 \cdot 1/3}{1/2} = 2/3$$

$\Rightarrow$ Wechseln lohnt sich

Unabhängigkeit von Ereignissen

- A, B unabhängig: Eintreten von A liefert keine Information über $P(B)$.
- Formal:

$$\boxed{P(A \mid B) = P(A)}$$

- Bei Unabhängigkeit ist äquivalent dazu:

$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

- Dann gilt:

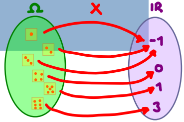
$$P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$$

- **Beispiel:** Werfen zweier Würfel:

$$\left.\begin{array}{l} A: \text{"'erster Würfel gleich 6"'} \\ B: \text{"'zweiter Würfel gleich 6"'} \end{array}\right\} \Rightarrow P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

$$= P(A)$$



$P(X=-1) = \frac{3}{6} = \frac{1}{2}$
$P(X=0) = \frac{1}{6}$

$P(X=0.5)=0$
$P(X\leq 0.5) = \frac{4}{6}$
$P(X \leq 1.1) = \frac{5}{6}$
$P(X \leq 2.99) = \frac{5}{6}$
$P(X \leq 3) = 1$
$F(x) = P(X \leq x)$

$$F(x) = \begin{cases} 0 & \text{für } x < -1 \\ \frac{1}{2} & \text{für } -1 \leq x < 0 \\ \frac{2}{3} & \text{für } 0 \leq x < 1 \\ \frac{5}{6} & \text{für } 1 \leq x < 3 \\ 1 & \text{sonst} \end{cases}$$

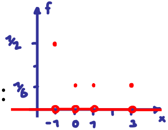
Zufallsvariablen und Verteilungen

- Beschreibung von Ereignissen durch reelle Zahlen
- Formal: Zufallsvariable ist Abbildung von Ereignisraum in reelle Zahlen:

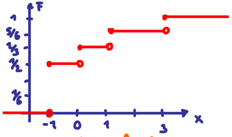
$$X: \ \Omega \to \mathbb{R}$$

- **Nach** Durchführung des Zufallsvorgangs:

Realisation: $x = X(\omega)$



$f(x) = P(X=x)$
Wahrscheinlichkeits-funktion



Verteilungsfunktion

- **Vor** Durchführung des Zufallsvorgangs:

Wertebereich: $X(\Omega) = \{x: \ x = X(\omega), \ \omega \in \Omega\}$

- **Beispiel:** Würfeln, X: Augenzahl, $X(\Omega) = \{1,2,\ldots,6\}$, $x = 4$ (z.B.)

$$P(X=4) = \frac{1}{6}, \quad P(X \leq 3) = \frac{3}{6} = \frac{1}{2}$$

1. Einführung
2. Deskriptive Statistik
3. W-Theorie
Kombinatorik
Zufall und Wahrscheinlichkeit
Zufallsvariablen und Verteilungen
Verteilungsparameter
4. Induktive Statistik
Quellen
Tabellen

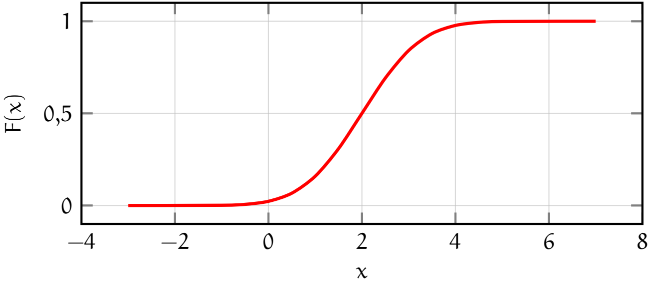
Verteilungsfunktion

- Zuweisung von Wahrscheinlichkeiten zu Realisationen
- Formal:

$$F(x) = P(X \leqq x)$$

- Eigenschaften der Verteilungsfunktion:
  - $F(x) \in [0; 1]$
  - Definitionsbereich: $\mathbb{R}$ mit $F(-\infty) = 0, \; F(\infty) = 1$
  - monoton wachsend, d.h. $x_1 < x_2 \Rightarrow F(x_1) \leqq F(x_2)$
  - Es gilt:

$$P(a < X \leqq b) = F(b) - F(a)$$



Beispiel einer Verteilungsfunktion

Diskrete Zufallsvariablen

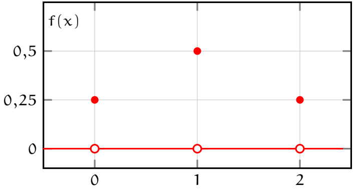
- X heißt diskret, wenn $X(\Omega) = \{x_1, x_2, \ldots\}$ endlich ist.
- Wahrscheinlichkeitsfunktion dann:

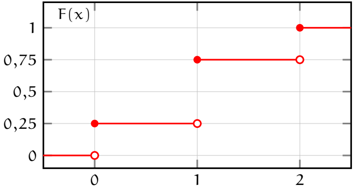
$$f(x) = P(X = x)$$

Beispiel: Münze 2 mal werfen; X: Anzahl "'Kopf"'

	(Z,Z)	$(Z,K),\ (K,Z)$	(K,K)
x_i	0	1	2
$f(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$F(x) = \begin{cases} 0, & \text{falls} \quad x < 0 \\ \frac{1}{4}, & \text{falls} \quad 0 \leqq x < 1 \\ \frac{3}{4}, & \text{falls} \quad 1 \leqq x < 2 \\ 1, & \text{falls} \quad x \geqq 2 \end{cases}$$





Binomialverteilung

Beispiel: 5 Versuche, Trefferw. pro Versuch: $p = 0{,}25$ (T(reffer), N(iete))

$X \mathrel{\widehat{=}}$ „Anzahl Treffer bei 5 Versuche"

$P(X=3) = P(\{(TTTNN),(TTNTN),\dots(NNTTT)\})$

$= \frac{5!}{3!\,2!}\cdot P(\{(TTTNN)\}) = \binom{5}{3}\, 0.25\cdot 0.25\cdot 0.25\cdot 0.75\cdot 0.75$

$= \binom{5}{3}\, 0.25^3\cdot 0.75^2 = 0{,}0879$

allgemein: p, Anzahl Vers.: n

$\boxed{P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}}$

- **Wiederholter** Zufallsvorgang
- n Durchführungen (jeweils unabhängig)
- Pro Durchführung: A oder $\bar{A}$ mit $P(A) = p$ ($\mathrel{\widehat{=}}$ Ziehen mit Zurücklegen)
- Schreibe:

$$X_i = \begin{cases} 1, & \text{falls } A \text{ bei } i\text{-ter Durchführung eintritt} \\ 0, & \text{falls } \bar{A} \text{ bei } i\text{-ter Durchführung eintritt} \end{cases}$$

- Dann gibt

$$X = \sum_{i=1}^{n} X_i$$

an, wie oft A eintritt.

- Gesucht: Wahrscheinlichkeitsfunktion von X

Binomialverteilung

- Herleitung:
  1) $P(X_i = 1) = P(A) = p, \; P(X_i = 0) = P(\bar{A}) = 1 - p$
  2) $\sum_{i=1}^{n} x_i = x$ entspricht "‚x mal Ereignis A und $n - x$ mal $\bar{A}$"'
     Wahrscheinlichkeit (bei Unabhängigkeit): $p^x \cdot (1-p)^{n-x}$
  3) Aber: Reihenfolge irrelevant! Anzahl Anordnungen: $\binom{n}{x}$

➠ **Wahrscheinlichkeitsfunktion der Binomialverteilung**:

$$f(x) = \begin{cases} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, & \text{falls } x \in \{0,1,\ldots,n\} \\ 0, & \text{sonst} \end{cases}$$

- Kurzschreibweise: $X \sim B(n; p)$
  X ist binomialverteilt mit Parametern n und p
- Tabellen zeigen meist $F(x)$
- für $f(x)$ gilt: $f(x) = F(x) - F(x-1)$

$X \sim B(n, 0.25)$, Tabelle der Binomialverteilung $F(x) = P(X \leqslant x)$

$x \setminus n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0.7500	0.5625	0.4219	0.3164	0.2373	0.1780	0.1335	0.1001	0.0751	0.0563	0.0422	0.0317	0.0238	0.0178	0.0134
1	1.0000	0.9375	0.8438	0.7383	0.6328	0.5339	0.4450	0.3671	0.3003	0.2440	0.1971	0.1584	0.1267	0.1010	0.0802
2		1.0000	0.9844	0.9492	0.8965	0.8306	0.7564	0.6786	0.6007	0.5256	0.4552	0.3907	0.3326	0.2811	0.2361
3			1.0000	0.9961	0.9844	0.9624	0.9295	0.8862	0.8343	0.7759	0.7133	0.6488	0.5843	0.5213	0.4613
4				1.0000	0.9990	0.9954	0.9871	0.9727	0.9511	0.9219	0.8854	0.8424	0.7940	0.7415	0.6865
5					1.0000	0.9998	0.9987	0.9958	0.9900	0.9803	0.9657	0.9456	0.9198	0.8883	0.8516
6						1.0000	0.9999	0.9996	0.9987	0.9965	0.9924	0.9858	0.9757	0.9617	0.9434
7							1.0000	1.0000	0.9999	0.9996	0.9988	0.9972	0.9944	0.9897	0.9827
8								1.0000	1.0000	1.0000	0.9999	0.9996	0.9990	0.9979	0.9958
9									1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992
10										1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
11											1.0000	1.0000	1.0000	1.0000	1.0000

$x \setminus n$	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
0	0.0100	0.0075	0.0056	0.0042	0.0032	0.0024	0.0018	0.0013	0.0010	0.0008	0.0006	0.0004	0.0003	0.0002	0.0002
1	0.0635	0.0501	0.0395	0.0310	0.0243	0.0190	0.0149	0.0116	0.0090	0.0070	0.0055	0.0042	0.0033	0.0025	0.0020
2	0.1971	0.1637	0.1353	0.1114	0.0913	0.0745	0.0607	0.0492	0.0398	0.0321	0.0258	0.0208	0.0166	0.0133	0.0106
3	0.4050	0.3530	0.3057	0.2631	0.2252	0.1917	0.1624	0.1370	0.1150	0.0962	0.0802	0.0666	0.0551	0.0455	0.0375
4	0.6302	0.5739	0.5187	0.4654	0.4149	0.3674	0.3235	0.2832	0.2467	0.2138	0.1844	0.1583	0.1354	0.1153	0.0979
5	0.8104	0.7653	0.7175	0.6678	0.6172	0.5666	0.5168	0.4685	0.4222	0.3783	0.3372	0.2990	0.2638	0.2317	0.2026
6	0.9205	0.8929	0.8610	0.8251	0.7858	0.7436	0.6994	0.6537	0.6074	0.5611	0.5154	0.4708	0.4279	0.3869	0.3481
7	0.9729	0.9598	0.9431	0.9226	0.8982	0.8701	0.8385	0.8037	0.7662	0.7265	0.6852	0.6427	0.5998	0.5568	0.5143
8	0.9925	0.9876	0.9807	0.9713	0.9591	0.9439	0.9254	0.9037	0.8787	0.8506	0.8196	0.7860	0.7502	0.7126	0.6736
9	0.9984	0.9969	0.9946	0.9911	0.9861	0.9794	0.9705	0.9592	0.9453	0.9287	0.9092	0.8868	0.8616	0.8337	0.8034
10	0.9997	0.9994	0.9988	0.9977	0.9961	0.9936	0.9900	0.9852	0.9787	0.9703	0.9599	0.9472	0.9321	0.9145	0.8943
11	1.0000	0.9999	0.9998	0.9995	0.9991	0.9983	0.9971	0.9954	0.9928	0.9893	0.9845	0.9784	0.9706	0.9610	0.9494
12	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9993	0.9988	0.9979	0.9966	0.9948	0.9922	0.9888	0.9842	0.9784
13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9997	0.9995	0.9991	0.9985	0.9976	0.9962	0.9944	0.9918
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9993	0.9989	0.9982	0.9973
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9997	0.9995	0.9992
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998
17		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Beispiel

Aus einem 32-er Kartenblatt wird 3-mal eine Karte mit Zurücklegen gezogen.
Wie wahrscheinlich ist es, 2-mal Herz zu ziehen?

$$X_i = \begin{cases} 1, & \text{falls } i\text{-te Karte Herz} \\ 0, & \text{sonst} \end{cases} \quad \Rightarrow \quad X_i \sim B(1; \tfrac{8}{32})$$

$$X = \sum_{i=1}^{n} X_i = X_1 + X_2 + X_3 \quad \Rightarrow \quad X \sim B(3; \tfrac{1}{4})$$

Mithilfe der Wahrscheinlichkeitsfunktion:

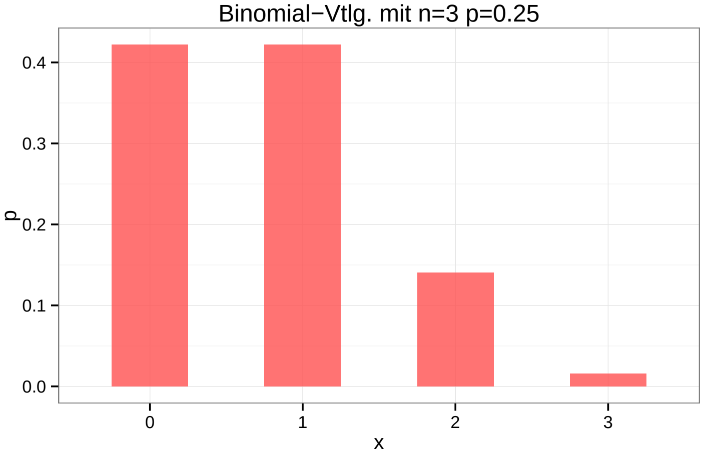
$$P(X = 2) = f(2) = \binom{3}{2} \cdot 0{,}25^2 \cdot 0{,}75^1 = 0{,}1406$$

Mithilfe der Tabelle ($n = 3$):

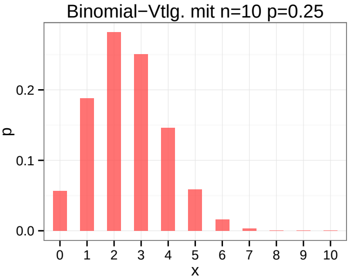
$$P(X = 2) = F(2) - F(1) = 0{,}9844 - 0{,}8438 = 0{,}1406$$

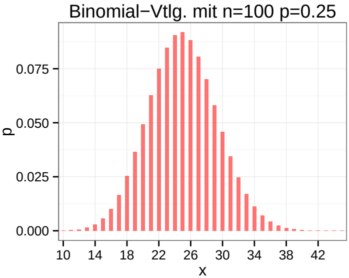
Binomialverteilung: Wahrscheinlichkeitsfunktion

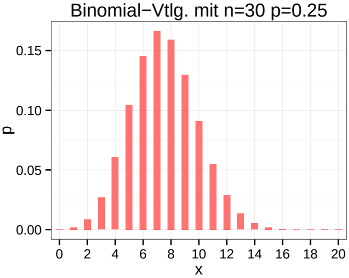
- $X \sim B(3, \frac{1}{4})$

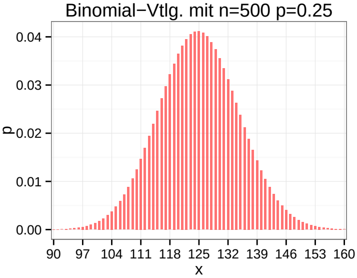


Binomialverteilung: Wahrscheinlichkeitsfunktion









Hypergeometrische Verteilung

- n-faches Ziehen **ohne** Zurücklegen aus N Objekten, davon M markiert.

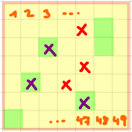
$$X = \text{Anzahl gezogener Objekte mit Markierung}$$

heißt hypergeometrisch verteilt mit den Parametern N, M, n.

- Kurzschreibweise: $X \sim \text{Hyp}(N; M; n)$
- Wahrscheinlichkeitsfunktion:

$$f(x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, & \text{falls } x \text{ möglich} \\ 0, & \text{sonst} \end{cases}$$

- Ist $n \leqq \frac{N}{20}$, so gilt: $\text{Hyp}(N; M; n) \approx B(n; \frac{M}{N})$



Lotto: 6 Richtige (unbekannt)

gesucht: Anzahl Möglichkeiten für „Dreier"

$$\frac{6\cdot 5\cdot 4}{3\cdot 2\cdot 1}\cdot\frac{43\cdot 42\cdot 41}{3\cdot 2\cdot 1}$$

$$=\binom{6}{3}\cdot\binom{43}{3} = 246\,820$$

Anzahl Mgl. 6 Kreuze beliebig zu setzen: $\binom{49}{6}$

$X \mathrel{\hat{=}}$ „Anzahl Treffer beim Lotto"

$$P(X=3) = \frac{\binom{6}{3}\cdot\binom{43}{3}}{\binom{49}{6}} \approx 0{,}0176$$

$$P(X=6) = \frac{\binom{6}{6}\cdot\binom{43}{0}}{\binom{49}{6}} \approx 7.151124e\text{-}08$$

LOTTO 6aus49 Gewinnquoten

Spieleinsatz: 57.265.834,00 €

Klasse	Anzahl Richtige	Gewinne	Quoten
1	6 Richtige + SZ	0 x	Jackpot
2	6 Richtige	5 x	412.478,30 €
3	5 Richtige + SZ	185 x	5.574,00 €
4	5 Richtige	1685 x	1.835,90 €
5	4 Richtige + SZ	9282 x	111,00 €
6	4 Richtige	83127 x	24,80 €
7	3 Richtige + SZ	137691 x	14,90 €
8	3 Richtige	1199722 x	7,70 €
9	2 Richtige + SZ	868797 x	5,00 €

Beispiel: Hypergeometrische Verteilung

- Aus einem 32-Kartenblatt wird 3-mal eine Karte ohne Zurücklegen gezogen.
- Wie wahrscheinlich ist es, 2-mal "'Herz"' zu ziehen?
- D.h.: $N = 32,\ M = 8,\ n = 3,\ x = 2.$

$$P(X = 2) = f(2) = \frac{\binom{8}{2}\binom{32-8}{3-2}}{\binom{32}{3}} = \frac{\binom{8}{2}\binom{24}{1}}{\binom{32}{3}} = \frac{\frac{8!}{2!\cdot 6!}\cdot 24}{\frac{32!}{3!\cdot 29!}}$$

$$= \frac{29!\cdot 8!\cdot 3!\cdot 24}{32!\cdot 6!\cdot 2!} = \frac{8\cdot 7\cdot 3\cdot 24}{32\cdot 31\cdot 30} = \frac{4032}{29760} = \frac{21}{155}$$

$$= 0{,}1355$$

Dabei wurde verwendet:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \text{und} \quad \binom{n}{1} = n.$$

Hypergeometrische Verteilung

Beispiel: x Treffer im Lotto 6 aus 49

- $X \sim \text{Hyp}(49, 6, 6)$

x	$P(X = x)$ (in %)
0	43.596498
1	41.301945
2	13.237803
3	1.765040
4	0.096862
5	0.001845
6	0.000007