Chapter 2: Probability Basic Concept (機率基本概念)

W. M. Song 桑慧敏
Tsing Hua Univ. 清華大學

2015.09.23

1. 第一個機率難題
2. 樣本, 事件空間, 機率函數 (S, $\ddot{E}$, P)
3. 三公設, 八定理, 量子力學上帝奧秘
4. 貝氏先生, 條件機率

Practice Problems

2011.10.18
(星期二)

車禍導致酵素飆高 酒測值破

她沒喝酒 卻有偽陽性反應

馬瑞君、馮惠宜／台中報導

卅二歲薛姓百貨公司專櫃小姐，騎機車下班途中與一名劉姓機車騎士對撞，被送醫後發現酒測值竟超過一毫克，已達「致死濃度」，因劉男和薛女同事都作證她沒飲酒，法醫研究所認為可能是車禍損傷造成身體酵素異常升高的「偽陽性酒精反應」，台中地院因而做出無罪判決。

去年八月五日，薛女騎車返家途中，與逆向行駛與劉姓機車騎士對撞。薛女被送至中國醫藥大學附設醫院急救，意識還算清醒，昏迷指數為十五分，但抽血檢測酒測值竟高達一‧七八六毫克，也就是達到「致死濃度」，會呈現神智不清、甚至昏迷情形。

薛女被依公共危險罪嫌起訴，但她否認喝酒，車禍當天上早班，上午十時半就上班，僅有喝牛奶和咖啡，就連中午、下午都沒有進食，直到下午六時就打卡下班，除上洗手間外，都沒有離開專櫃。

同專櫃的陳姓彩妝師也作證，當天因為很忙，看到薛女到下班前都沒有去用餐，因專櫃不能帶飲料上班，也沒有看到她喝飲料或水，看到她下班時精神很正常。劉姓機車騎士也作證，車禍發生時，他和朋友在現場都沒有聞到薛女身上有酒味。

負責照顧薛女的急診室陳姓醫師作證，薛女被送入急診室時意識清楚、還可以溝通，問一些狀況都能回答；醫護人員也沒有聞到薛女嘔吐物或口鼻有酒味。

台中地院檢送法醫研究所鑑定
答案是以酵素法進行酒測時，可
受傷或疾病因素造成血液中異常
酸鹽和乳酸去酶會產生干擾，使
偽陽性，這種情況在臨床上並不
必須將此種偽陽性的可能列入必
薛女車禍後兩小時才抽血，若
正確，她車禍時的酒測值早已超
度，不可能正常對話或提供電話
個人清醒，更不可能平衡機車駕
中山醫學大學檢驗暨生物學系
江指出，某些物質、食物或代謝
能產生同樣的代謝物干擾檢驗結
第一線搞錯檢體，的確有一些物
會干擾檢驗結果。

2011.10.18 中國時報: 她沒喝酒, 卻有偽陽性反應. 請用符號定義何謂 “偽陽性” (false positive)

Probability (機率)

Probability is a

- real number (實數)
- variable (變數)
- function (函數)
- Others (其他)
- Ans: function (函數)

Probability function, P

Domain and co-domain

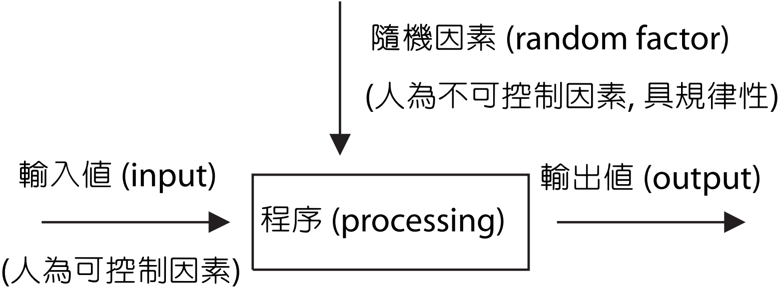
- "Domain" (定義域): $\mathcal{D} = \{\text{all events}\}$
- all subsets of a sample space, S
- "Co-Domain" (對應域):
- $[0, 1]$

- 投擲一個公正的錢幣 1 次, 如何描述其機率空間 (S, $\ddot{E}$, P)?
- S: sample space; $\ddot{E}$: event space; P: probability function.

機率基本概念 (三輪車調)

- 分賭金, 算機率, 一點五個世紀難題;
- 樣本空間, 事件空間, 機率函數, 有定義;
- 三公設, 八定理, 量子力學上帝奧秘;
- 貝氏先生, 條件機率, 三八加一線, 全備齊;
- 討論: 歌詞含意

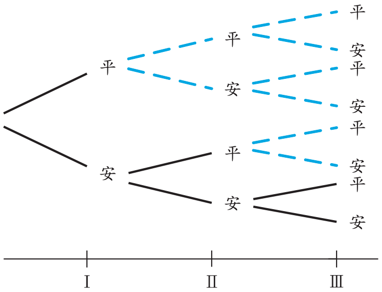
隨機實驗



- 討論: 隨機實驗與確定性實驗有何不同?

分賭金, 算機率, 一點五個世紀難題

分獎金問題 (修改於 Pacioli 中之賭金的分配問題):
帥安與聖平兩個實力相當的桌球選手比賽桌球, 誰先拿到21 分就可得獎金一萬元。比賽到中場時發生地震必須停止比賽, 此時帥安拿到 18 分, 聖平拿到 20 分。如何公平的分獎金?



- 《算數, 幾何, 及比例性質之摘要》, Pacioli (1494) 在意大利出版的第一本有關機率的書, 書中提到分賭金問題, 一點五個世紀後才得解。

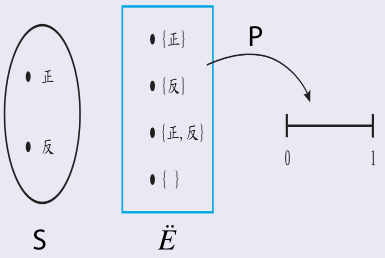
機率空間 (S, $\ddot{E}$, P)

- 樣本點 (sample point): 任一個隨機實驗中可能出現的每一個結果 (outcome), 稱為樣本點, 以小寫 s 表示之。
- 樣本空間 (sample space): 任一個隨機實驗中, 包含所有樣本點的集合稱為樣本空間, 以英文大寫字母 S 表示之。
- 事件 (event): 樣本空間的部分集合稱為事件。
- 事件空間 (event space) $\ddot{E}$: 我們有興趣的所有事件所成的集合稱為事件空間。事件以 E 表示, 上方兩點表示空間
- 機率 (P): 是一個函數, 將隨機實驗中的任何一個事件對應到介於0與1中的某一實數。

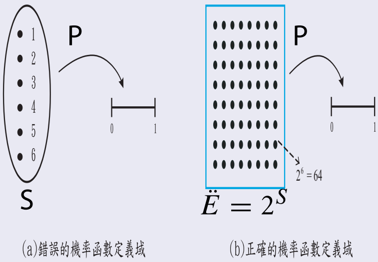
- 討論: S, $\ddot{E}$, P

機率空間 (S, $\ddot{E}$, P)

Flip a fair coin once



Flip a fair die once



- 機率函數 P 的定義域是 $\ddot{E}$, 不是 S

「上帝是不擲骰子的」: 愛因斯坦錯了嗎?

- 玻爾 (N. Bohr):「波函數的統計詮釋, 獲得 1929 年 Nobel 物理獎。

  在任一時刻, 粒子的所在位置與動量以機率的方式來描述比以確定性的方式來描述更能吻合實驗的結果, 同時在某些狀況下, 動量服從不連續的機率分配 。

- 愛因斯坦在 1926 年 12 月寫當時的物理學家波恩 (M. Born) 信中說到:

  「量子力學是令人讚嘆的, 但是有一個內在的聲音告訴我, 這還不是真正的貨色。這個理論有很大的貢獻, 但是它並不使我們更接近上帝的奧秘一些。無論如何, 我相信上帝是不擲骰子的」。

機率三公設

公設1: $\mathsf{P}(\mathsf{S}) = 1$

公設2: 對任一事件 $E \in \ddot{E}, 0 \leq \mathsf{P}(E) \leq 1$

公設3: 如果事件 E_1 與 E_2 互斥 (i.e., $E_1 \cap E_2 = \varnothing$), 則

$$\mathsf{P}(E_1 \cup E_2) = \mathsf{P}(E_1) + \mathsf{P}(E_2)$$

- 討論: 機率三公設含意

八個基本機率定理

- Thm 2.1. $\mathrm{P}(\phi) = 0$
- Thm 2.2. 如果 $E_1 \subset E_2$, 則 $\mathrm{P}(E_1) \leq \mathrm{P}(E_2)$
- Thm 2.3. 對於任何 n 個互斥事件 $E_1, \ldots, E_n$, $\mathrm{P}(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \mathrm{P}(E_i)$
- Thm 2.4. 如果 n 個互斥事件 $E_1, E_2, \ldots, E_n$ 發生機率均等, 且已知 $E_1 \cup E_2 \cup \ldots \cup E_n = \mathrm{S}$, 則 $\mathrm{P}(E_i) = 1/n, i = 1, 2, \ldots, n$
- Thm 2.5. 假設一個隨機實驗的樣本空間 S 包含有限個樣本點, 且每一個樣本點出現機會均等。則 $\mathrm{P}(E) = m/n$, 其中 m 為事件 E 中元素個數, n 為 S 中所有樣本點個數。
- Thm 2.6.對於任何兩個事件 E_1 和 E_2, $\mathrm{P}(E_1 \cup E_2) = \mathrm{P}(E_1) + \mathrm{P}(E_2) - \mathrm{P}(E_1 \cap E_2)$
- Thm 2.7.對於任何三個事件 E_1, E_2, 和 E_3, $\mathrm{P}(E_1 \cup E_2 \cup E_3) =$ $\mathrm{P}(E_1) + \mathrm{P}(E_2) + \mathrm{P}(E_3) - \mathrm{P}(E_1 \cap E_2) - \mathrm{P}(E_1 \cap E_3) - \mathrm{P}(E_2 \cap E_3) + \mathrm{P}(E_1 \cap E_2 \cap E_3)$
- Thm 8. $\mathrm{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^n \mathrm{P}(E_i) - \sum_{i<j} \mathrm{P}(E_i \cap E_j) +$ $\sum_{i<j<k} \mathrm{P}(E_i \cap E_j \cap E_k) \cdots + (-1)^{n+1} \mathrm{P}(E_1 \cap E_2 \cap \cdots \cap E_n)$

2011.10.18 中國時報

2011.10.18
(星期二)

車禍導致酵素飆高 酒測值破

她沒喝酒 卻有偽陽性反應

馬瑞君、馮惠宜／台中報導

卅二歲薛姓百貨公司專櫃小姐，騎機車下班途中與一名劉姓機車騎士對撞，被送醫後發現酒測值竟超過一毫克，已達「致死濃度」，因劉男和薛女同事都作證她沒飲酒，法醫研究所認為可能是車禍損傷造成身體酵素異常升高的「偽陽性酒精反應」，台中地院因而做出無罪判決。

去年八月五日，薛女騎車返家途中，與逆向行駛與劉姓機車騎士對撞。薛女被送至中國醫藥大學附設醫院急救，意識還算清醒，昏迷指數為十五分，但抽血檢測酒測值竟高達一．七八六毫克，也就是達到「致死濃度」，會呈現神智不清、甚至昏迷情形。

薛女被依公共危險罪嫌起訴，但她否認喝酒，車禍當天上早班，上午十時半就上班，僅有喝牛奶和咖啡，就連中午、下午都沒有進食，直到下午六時就打卡下班，除上洗手間外，都沒有離開專櫃。

同專櫃的陳姓彩妝師也作證，當天因為很忙，看到薛女到下班前都沒有去用餐，因專櫃不能帶飲料上班，也沒有看到她喝飲料或水，看到她下班時精神很正常。劉姓機車騎士也作證，車禍發生時，他和朋友在現場都沒有聞到薛女身上有酒味。

負責照顧薛女的急診室陳姓醫師作證，薛女被送入急診室時意識清楚，還可以溝通，問一些狀況都能回答；醫護人員也沒有聞到薛女嘔吐物或口鼻有酒味。

台中地院檢送法醫研究所鑑定……答案是以酵素法進行酒測時，可……受傷或疾病因素造成血液中異常……酸鹽和乳酸去酶會產生干擾，使……偽陽性，這種情況在臨床上並不……必須將此種偽陽性的可能列入必……

薛女車禍後兩小時才抽血，若……正確，她車禍時的酒測值早已超……度，不可能正常對話或提供電話……個人清醒，更不可能平衡機車駕……

中山醫學大學檢驗暨生物學系……江指出，某些物質、食物或代謝……能產生同樣的代謝物干擾檢驗結……第一線搞錯檢體，的確有一些物……會干擾檢驗結果。

Q: What is 偽陽性 (false positive)?

Conditional Probability

$\mathsf{P}(B|A)$

The conditional probability of an event B given an event A, denoted by $\mathsf{P}(B|A)$, is

$$\mathsf{P}(B|A) = \begin{cases} \dfrac{\mathsf{P}(A \cap B)}{\mathsf{P}(A)}, & \mathsf{P}(A) > 0 \\ 0, & \mathsf{P}(A) = 0 \end{cases}$$

- Q: Why is the above definition appropriate?
- A: The definition does not violate the 3 axioms

- $\mathsf{P}(\mathsf{S}|A) = 1$ (need to prove)
- 對任意事件 E, $0 \leq \mathsf{P}(E|A) \leq 1$ (need to prove)
- 如果事件 B 與 C 互斥, 則 $\mathsf{P}(B \cup C|A) = \mathsf{P}(B|A) + \mathsf{P}(C|A)$ (need to prove)

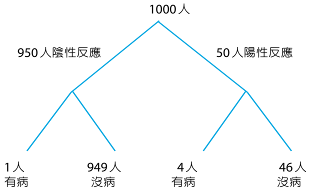
Conditional 8 Theorems

- Thm 2.9. $\mathsf{P}(\phi|A) = 0$
- Thm 2.10. If $E_1 \subset E_2$, then $\mathsf{P}(E_1|A) \leq \mathsf{P}(E_2|A)$
- Thm 2.11. If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then $\mathsf{P}(\bigcup_{i=1}^{n} E_i|A) = \sum_{i=1}^{n} \mathsf{P}(E_i|A)$
- Thm 2.12. For any two events E_1 and E_2, $\mathsf{P}(E_1 \cup E_2|A) = \mathsf{P}(E_1|A) + \mathsf{P}(E_2|A) - \mathsf{P}(E_1 \cap E_2|A)$
- Thm 2.13. For any three events E_1, E_2 和 E_3, $\mathsf{P}(E_1 \cup E_2 \cup E_3|A) = \mathsf{P}(E_1|A) + \mathsf{P}(E_2|A) + \mathsf{P}(E_3|A)$ $- \mathsf{P}(E_1 \cap E_2|A) - \mathsf{P}(E_1 \cap E_3|A) - \mathsf{P}(E_2 \cap E_3|A) + \mathsf{P}(E_1 \cap E_2 \cap E_3|A)$
- Thm 2.14. For any n events $E_1, E_2, \ldots, E_n$, $\mathsf{P}(\cup_{i=1}^{n} E_i|A) = \sum_{i=1}^{n} \mathsf{P}(E_i|A) - \sum_{i<j} \mathsf{P}(E_i \cap E_j|A)$ $+ \sum_{i<j<k} \mathsf{P}(E_i \cap E_j \cap E_k|A) \ldots + (-1)^{n+1}\mathsf{P}(E_1 \cap E_2 \cap \cdots \cap E_n|A)$
- Comparing “Conditional Probabilities Theorems” and the corresponding Probabilities Theorems

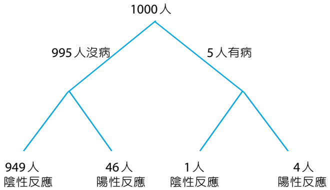
Example

假設 A 區有 1000 個人口, 其中有 5 人患有某種疾病。若對此 1000 經過測試後, 發現有 950 人呈現陰性反應 (顯示沒病), 50 人呈現陽性反應 (顯示有病)。事實上, 呈現陰性反應的 950 人中有 1 人確實有病, 而呈現陽性反應的 50 人中 4 人確實有病。令 A 為有病的事件; B 為陽性反應的事件; A^{C} 為無病的事件; B^{C} 為陰性反應的事件.

1. 一般人患有此病的機率為何 $\mathrm{P}(A)$?
2. 檢驗結果是陽性的機率為何 $\mathrm{P}(B)$?
3. 「偽陰性」(有病卻被檢驗出陰性反應) 的機率為何 $\mathrm{P}(B^{\mathsf{C}}|A)$?
4. 「偽陽性」(沒病卻被檢驗出陽性反應) 的機率為何 $\mathrm{P}(B|A^{\mathsf{C}})$?



(a) 表示法一



(b) 表示法二

- Advantage of using notations and plots.

Multiplicative Rule

Theorem 2.15

For every two events A and B, $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

Theorem 2.16

For any n events $E_i, i = 1, 2, \cdots, n$,
$P(\bigcap_{i=1}^{n} E_i) = P(E_1)P(E_2|E_1)\cdots P(E_n|\bigcap_{i=1}^{n-1} E_i)$

$n = 2$, $P(E_1 \cap E_2) = P(E_1)P(E_2|E_1)$ $\underline{Theorem 2.15}$

$n = k$, $P(\bigcap_{i=1}^{k} E_i) = P(E_1)P(E_2|E_1)\cdots P(E_k|\bigcap_{i=1}^{k-1} E_i)$

$$
\begin{aligned}
n = k+1,\ P(\textstyle\bigcap_{i=1}^{k+1} E_i) &= P[(\textstyle\bigcap_{i=1}^{k} E_i) \cap E_{k+1}] && \underline{(Q1)}\\
&= P(\textstyle\bigcap_{i=1}^{k} E_i)P(E_{k+1}|\textstyle\bigcap_{i=1}^{k} E_i) && \underline{(Q2)}\\
&= P(E_1)P(E_2|E_1)\cdots P(E_k|\textstyle\bigcap_{i=1}^{k-1} E_i)P(E_{k+1}|\textstyle\bigcap_{i=1}^{k} E_i)
\end{aligned}
$$

- What method (induction?) is used for proving of Them 2.16?
- Mathematical induction. (Q1): Intersection; (Q2): Theorem 2.15

總機率定理

如果事件 $E_1, E_2, \ldots, E_n$ 是樣本空間的一個分割 (partition), 則任一事件 A 可寫成
$\mathrm{P}(A) = \sum_{i=1}^{n} \mathrm{P}(A \cap E_i)$

貝氏定理

如果事件 $E_1, E_2, \ldots, E_n$是樣本空間的一個分割, if $\mathrm{P}(B) > 0$,

$$\mathrm{P}(E_1|B) = \frac{\mathrm{P}(E_1 \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(E_1)\mathrm{P}(B|E_1)}{\sum_{i=1}^{n} \mathrm{P}(E_i)\mathrm{P}(B|E_i)}$$

統計獨立: independence

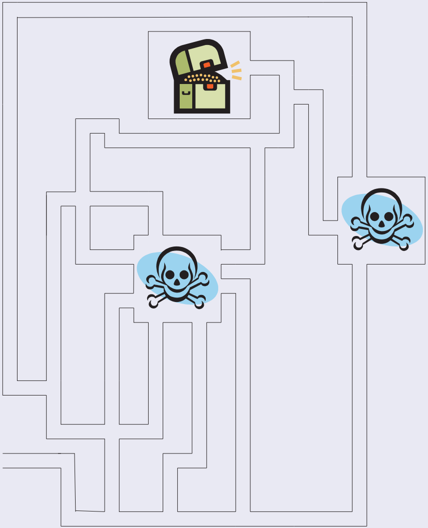
Two events A and B are independent if and only if one of the following condition holds.

- 1. $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$
- 2. $\mathrm{P}(A|B) = \mathrm{P}(A)$
- 3. $\mathrm{P}(B|A) = \mathrm{P}(B)$

Ex. 三個臭皮匠勝過一個諸葛亮: 令 A, B, C 分別為三個臭皮匠獨力解決問題的事件, 則 P(A), P(B), P(C) 分別為三個臭皮匠獨力解決問題的機率。假設臭皮匠能互不干擾的解決問題。若三個臭皮匠合力解決問題, 則問題被解決的的機率為

$$\mathrm{P}(A \cup B \cup C) = 1 - [1 - \mathrm{P}(A)][(1 - \mathrm{P}(B)][1 - \mathrm{P}(C)] = 1 - (1 - 0.45)(1 - 0.55)(1 - 0.60) = 0.901$$

Ex. 2.22. 迷宮盜墓



Define Notation:

- Event A: Succeed
- E_i: correct selection in the $i-th$ intersection
- There are 6 inter. with 2, 3, 2, 3, 2, 2 choices

$$
\begin{aligned}
&\mathsf{P}(A) \\
&\quad = \mathsf{P}(E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5 \cap E_6) \\
&\quad = \mathsf{P}(E_1)\mathsf{P}(E_2|E_1)\mathsf{P}(E_3|E_1 \cap \\
&\qquad E_2) \cdots \mathsf{P}(E_6| \textstyle\bigcap_{i=1}^{5} E_i) \\
&\quad = (\tfrac{1}{2})(\tfrac{1}{3})(\tfrac{1}{2})(\tfrac{1}{3})\tfrac{1}{2})(\tfrac{1}{2}) \\
&\quad = 1/144
\end{aligned}
$$

Ex. Monty Hall Problem

有三個門讓你選, 三個門中只有一個門後有大獎, 其他兩個門後沒有獎。你選定後, 主持人打開其中一個你沒選中也沒有獎的門。現在再讓你作一次選擇, 你可以選擇換門或不換門。如果你選擇換門, 那得獎的機率是多少?

Analytical Appro. I

- $\mathsf{S} = \{(i, j), i = 1, 2, 3; j = 1, 2, 3\}$
- E =「參與者最後中獎」的事件
- $= \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$, 一共有 6 個樣本點
- 計算: P(E) $= 6/9 = 2/3$ (Thm 2.5)

Analytical Appro. II

- $\mathsf{S} = \{1, 2, 3\}$
- F =「第一次就選中有獎的門」的事件
- $\mathsf{P}(E) = \mathsf{P}(E \mid F)\mathsf{P}(F) + \mathsf{P}(E \mid F^{\mathsf{C}})\mathsf{P}(F^{\mathsf{C}}) = (0)(1/3) + (1)(2/3) = 2/3$

- 參考 "桑慧敏, 機率與推論統計學原理. p. 55, 68"