Axiomatic Definition of Probability

Let Ω be a sample space and let $\mathcal{B}$ be an event space (of subsets of Ω which forms a σ-field)

A set function $P : \mathcal{B} \to \mathbb{R}$ is said to be a probability function if it satisfies the following three axioms :

P1 : $P(E) \geq 0 \ \forall \ E \in \mathcal{B}$

ie prob is always non-negative.

P2 : $P(\Omega) = 1$

P3 : Let $\{E_i\}$ be a sequence of pairwise disjoint sets in $\mathcal{B}$,

then $$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

This is countable additivity.

$(\Omega, \mathcal{B}, P)$ is called a probability space.

<u>Some Properties</u> of <u>the Probability Function</u>:

1. $P(\phi) = 0$

<u>Pf</u>. In P3, let us take $E_1 = \Omega$ & $E_2 = E_3 = \cdots = \phi$. Then

$$P(\Omega) = P(\Omega) + P(\phi) + P(\phi) + \cdots$$

$$\Rightarrow 1 = 1 + P(\phi) + P(\phi) + \cdots$$

$$\Rightarrow P(\phi) = 0$$

2. For any finite collection $\{E_1, \ldots, E_n\}$ of pairwise disjoint events in $\mathcal{B}$,

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

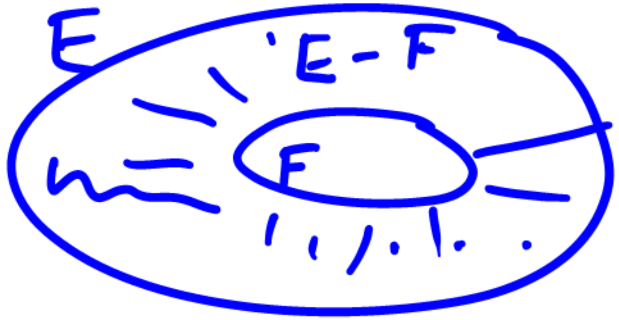
Pf. In P3, let us take $E_{n+1} = E_{n+2} = \cdots = \phi$.

Then $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) + P(\phi) + P(\phi) + \cdots$

$\Rightarrow P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$

So prob fn is finitely additive.

3. Let $\boxed{F \subset E}$.



$$E = F \cup (E - F)$$

$$P(E) = P(F) + P(E - F)$$

$$\Rightarrow \quad P(F) \le P(E) \quad \cdots \quad (*)$$

& also $P(E - F) = P(E) - P(F)$

ie prob is a monotonic function

4. Since $\phi \subset E \subset \Omega$

for any $E \in \mathcal{B}$

$$\Rightarrow P(\phi) \leq P(E) \leq P(\Omega)$$

$$\Rightarrow 0 \leq P(E) \leq 1$$

for any event.

5. $\Omega = A \cup A^c$ for any $A \in \mathcal{B}$

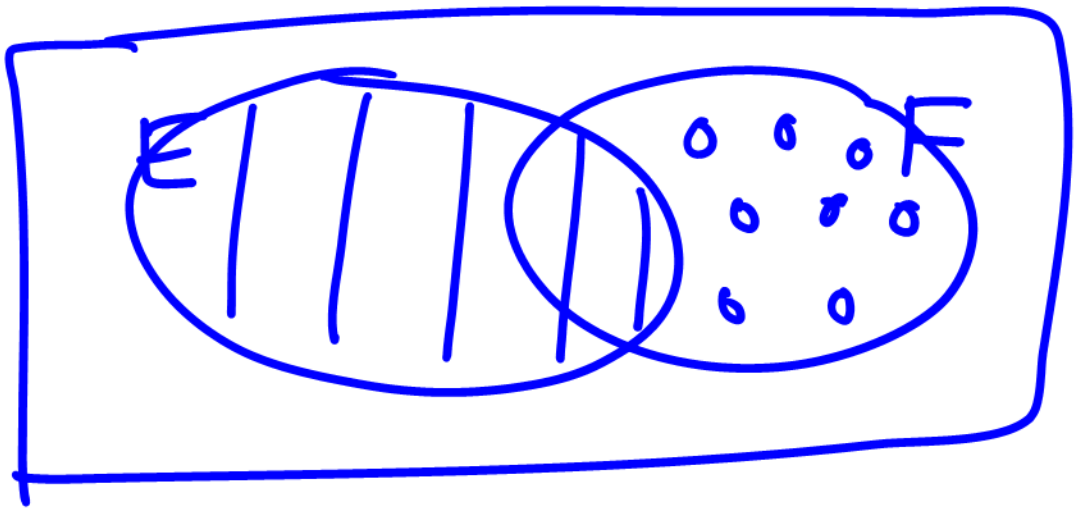
$$\Rightarrow P(\Omega) = P(A) + P(A^c)$$

$$\Rightarrow 1 = P(A) + P(A^c) \Rightarrow P(A^c) = 1 - P(A)$$

6. **Addition Rule**: Let $E, F \in \mathcal{B}$.

Then $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Pf.



$$E \cup F = E \cup (F - (E \cap F))$$

So $P(E \cup F) = P(E) + P(F - (E \cap F))$

$$= P(E) + P(F) - P(E \cap F).$$

7. <u>General Addition Rule</u> :

For any events $E_1, \ldots, E_n \in \mathcal{B}$,

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) - \sum_{i<j}\sum P(E_i \cap E_j)$$

$$+ \sum_{i<j<k}\sum\sum P(E_i \cap E_j \cap E_k)$$

$$- \cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} E_i\right)$$

<u>Proof</u> (by induction)

For $n=1$ $\rightarrow$ trivial

For $n=2$, we have proved the addition rule.

Let us assume it to be true for $n=r$.

Take $n = r+1$

$$P\left(\bigcup_{i=1}^{r+1} E_i\right) = P\left(\underbrace{\left(\bigcup_{i=1}^{r} E_i\right)} \cup \underbrace{E_{r+1}}\right)$$

$$= \underbrace{P\left(\bigcup_{i=1}^{r} E_i\right)} + P\left(E_{r+1}\right) - P\left(\left(\bigcup_{i=1}^{r} E_i\right) \cap E_{r+1}\right)$$

$$= \left\{ \sum_{i=1}^{r} P(E_i) - \sum^{r}\sum^{r}_{i<j} P\left(E_i \cap E_j\right) \right.$$

$$+\sum^{r}\sum^{r}\sum^{r}_{i<j<k} P(E_i \cap E_j \cap E_k)$$

$$- \cdots + (-1)^{r+1} P\left(\bigcap_{i=1}^{r} E_i\right)\Bigg] + \underline{P(E_{r+1})}$$

$$\underbrace{-P\left(\bigcup_{i=1}^{r} (E_i \cap E_{r+1})\right)}$$

$$= \sum_{i=1}^{r+1} P(E_i) - \sum^{r}\sum^{r}_{i<j} P(E_i \cap E_j)$$

$$+\sum_{i<j<k}^{r}\sum^{r}\sum^{r} P(E_i \cap E_j \cap E_k) - \cdots$$

$$+(-1)^{r+1} P\left(\bigcap_{i=1}^{r} E_i\right) - \left[\sum_{i=1}^{r} P(E_i \cap E_{r+1})\right.$$

$$-\sum_{i<j}^{r}\sum^{r} P(E_i \cap E_j \cap E_{r+1})$$

$$+\sum_{i<j<k}^{r}\sum^{r}\sum^{r} P(E_i \cap E_j \cap E_k \cap E_{r+1})$$

$$- \cdots + (-1)^{r+1} P\left(\bigcap_{i=1}^{r+1} E_i\right)\Bigg]$$

$$= \sum_{i=1}^{r+1} P(E_i) - \sum_{i<j}^{r+1}\sum^{r+1} P(E_i \cap E_j)$$

$$+ \sum_{i<j<k}^{r+1}\sum^{r+1}\sum^{r+1} P(E_i \cap E_j \cap E_k)$$

$$- \cdots + (-1)^{r+2} P\left(\bigcap_{i=1}^{r+1} E_i\right).$$

This proves the result.

8. For $E_1, \ldots, E_n \in \mathcal{B}$,

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

(Sub-additivity of Prob. fn)

9. For any countable sequence $\{E_i\}$ in $\mathcal{B}$,

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} P(E_i).$$

9. <u>Bonferroni Inequality</u> : For any events $E_1, \ldots, E_n \in \mathcal{B}$

$$\sum_{i=1}^{n} P(E_i) - \sum_{i<j}^{n}\sum^{n} P(E_i \cap E_j) \leq P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

(*) Prove above results yourself.

10. Boole's Inequality: Let $\{E_n\} \in \mathcal{B}$.

Then $P\left(\bigcap_{i=1}^{\infty} E_i\right) \geq 1 - \sum_{i=1}^{\infty} P(E_i^c)$

Pf. $P\left(\bigcap_{i=1}^{\infty} E_i\right) = 1 - P\left(\left(\bigcap_{i=1}^{\infty} E_i\right)^c\right)$

$= 1 - P\left(\bigcup_{i=1}^{\infty} E_i^c\right)$

$\geq 1 - \sum_{i=1}^{\infty} P(E_i^c)$.

Example: Birthday problem

Suppose there are n students in a class. Assume $n \leq 365$ and no student has a birthday on 29^{th} feb.

What is the prob that no two students share a common b'day.

$A \rightarrow$ No two students share a common b'day

$$P(A) = \frac{{}^{365}P_n}{(365)^n}$$ assuming all dates to be equally likely for birth

$$= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right)$$

$A^c \rightarrow$ at least two persons share a b'day

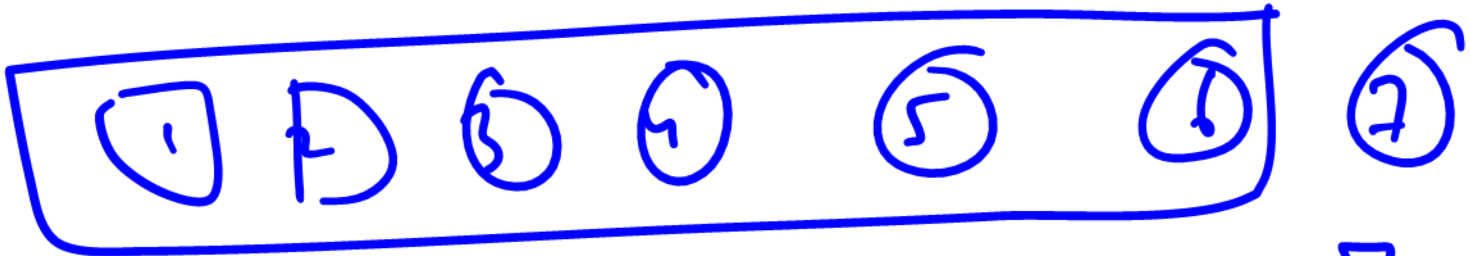
$$P(A^c) = 1 - P(A)$$

n	$P(A)$	$P(A^c)$
10	0.871	0.129
20	0.589	0.411
23	0.493	0.507
30	0.294	0.706
50	0.030	0.970
60	0.006	0.994

(*) Ex : What is the prob of 3 persons sharing a b'day, ... and so on ?

2. Let 7 balls be placed randomly in 7 cells. What is the prob that exactly one cell remains empty?

Sol^n



One empty cell can be selected in 7 ways. Now one cell will have two balls. This cell can be selected in $^6C_1 = 6$ ways.

Two balls can be selected out of 7 in ${}^7C_2$ ways. Remaining 5 balls can be placed in 5 cells in 5! ways.

So the reqd prob. $= \dfrac{7 \times 6 \times {}^7C_2 \times 5!}{7^7}$

$= \dfrac{2160}{16807} \approx 0.1285.$

3. A pair of fair dice is tossed. If an odd sum or a sum of 4 appears, we stop. Else the toss is continued. What is the prob that 4 appears?

Solⁿ Let $E \to$ odd sum appears

There are total 18 elements in E.

Let $F \to$ sum is 4, $F = \{(1,3),(3,1),(2,2)\}$

So $P(4 \text{ appears}) = P(4 \text{ appears in } I^{st} \text{ trial})$

$+ P(4 \text{ appears on } 2^{nd} \text{ trial}) + \cdots$

$$= \frac{1}{12} + \boxed{\frac{5}{12} \times \frac{1}{12}} + \left(\frac{5}{12}\right)^2 \times \frac{1}{12} + \cdots$$

$$= \frac{1}{7}.$$

Conditional Probability

Let $(\Omega, \mathcal{B}, P)$ be a prob. space and let B be any event with $P(B) > 0$.

For any event $A \in \mathcal{B}$, we define the conditional probability of A given that B has already occurred as

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

$$\Rightarrow P(A \cap B) = P(B)\, P(A|B)$$

If we assume $P(A) > 0$, then we can

also have

$$P(A \cap B) = P(A)\, P(B|A)$$

Multiplication Rule of Probability

General Multiplication Rule

Let $A_1, A_2, \ldots, A_n \in \mathcal{Q}$ with

$P\left(\bigcap_{i=1}^{n} A_i\right) > 0$. Then

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \cdots P\left(A_n \,\Big|\, \bigcap_{i=1}^{n-1} A_i\right).$$

Proof : $\star$

Let $B_1, B_2, \ldots, B_n$ be pairwise disjoint events with $B = \bigcup_{i=1}^{n} B_i$

& $P(B_j) > 0$, $j = 1, \ldots, n$

Then $P(A \cap B) = \sum_{j=1}^{n} P(A \mid B_j)\, P(B_j)$

$$A \cap B = A \cap \left(\bigcup_{j=1}^{n} B_j \right) = \bigcup_{j=1}^{n} (A \cap B_j)$$

$$\underline{P(A \cap B)} = \sum_{j=1}^{n} P(A \cap B_j)$$

$$= \sum_{j=1}^{n} P(A \mid B_j)\, P(B_j)$$

When $B = \Omega$, ie, $B_1, \ldots B_n$ are exhaustive events, then

$$P(A) = \sum_{j=1}^{n} P(A \mid B_j) P(B_j)$$

This Theorem of Total Probability.

<u>Example</u>: A computer manufacturer procures IC's from suppliers B_1, B_2, B_3

with 40% from B_1, 30% from B_2 & 30% from B_3. Suppose 1% of supply from B_1 is defective, 5% from B_2, 10% from B_3 is defective. What is the prob that a randomly selected IC from the stock is defective.

$A \rightarrow$ IC is defective

$$P(A) = \sum_{j=1}^{3} P(A \mid B_j)\, P(B_j)$$

$= 0.01 \times 0.4 + 0.05 \times 0.3$

$+ 0.1 \times 0.3 = 0.049$

<u>Bayes Theorem</u> (Thomas Bayes → 1763)

Let $B_1, B_2, \ldots, B_n$ be pairwise disjoint exhaustive events ($\cup B_i = \Omega$) with $P(B_j) > 0 \ \forall j$. Let $A \in \mathcal{B}$, with $P(A) > 0$.

Then

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)\, P(B_j)}$$

__Pf.__ $P(B_i \mid A) = \dfrac{P(B_i \cap A)}{P(A)}$

$$= \frac{P(A \mid B_i)\, P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)\, P(B_j)}.$$

Example (Defective IC's): Suppose a randomly selected IC is found to be defective. What is the prob that it was supplied by B_1 (or B_2 or B_3)

$$P(B_1|A) = \frac{P(A|B_1)\,P(B_1)}{P(A)} = \frac{0.01 \times 0.4}{0.049}$$

$$= \frac{4}{49}, \quad P(B_2|A) = \frac{15}{49}, \quad P(B_3|A) = \frac{30}{49}$$

Independence of Events

If $P(A|B) = P(A)$

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

$$\Rightarrow P(B|A) = P(B)$$

Events A and B are said to be statistically independent if

$$P(A \cap B) = P(A)\,P(B)$$

A, B, C are independent if

$P(A \cap B) = P(A)P(B)$, $P(B \cap C) = P(B)P(C)$

$P(C \cap A) = P(C)P(A)$, $P(A \cap B \cap C) = P(A)P(B)P(C)$

$A_1, \ldots, A_n$ are independent if

$P(A_i \cap A_j) = P(A_i)P(A_j) \quad \forall\, i < j$

$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$

$\forall\, i < j < k$

$$\cdots\quad P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

$(2^n - n - 1$ conditions$)$

<u>Example</u>: Pairwise independence does not imply mutual independence.

Suppose two fair dice are tossed

$A \rightarrow$ odd no on first dice

$B \rightarrow$ odd no on second dice

$C \to$ sum is odd.

$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(C) = \frac{1}{2}$

$P(A \cap B) = \frac{1}{4}$, $P(A \cap C) = \frac{1}{4}$, $P(B \cap C) = \frac{1}{4}$

So A, B, C are pairwise independent

$P(A \cap B \cap C) = 0 \neq P(A)\,P(B)\,P(C)$.

So A, B, C are not independent.

<u>Example</u>: Six cards are drawn with replacement from a deck of 52 cards. What is the prob that each of the four suits will be represented at least once among the six cards?

<u>Sol^n</u>: $A \rightarrow$ all suits appear at least once

$A^c \rightarrow$ at least one suit does not appear

$$A^c = \bigcup_{i=1}^{4} B_i$$

$B_1 \to$ spades do not appear, $B_2 \to$ hearts don't

$B_3 \to$ diamonds do not, .. $B_4 \to$ clubs don't ...

$P(B_1) = P(\text{none of six cards is a spade})$

$$= \left(\frac{3}{4}\right)^6 = P(B_2) = P(B_3) = P(B_4)$$

$$P(B_1 \cap B_2) = \left(\frac{1}{2}\right)^6 = P(B_i \cap B_j) \ \forall\ i<j$$

$$P(B_i \cap B_j \cap B_k) = \left(\frac{1}{4}\right)^6, \ i<j<k$$

$$P\left(\bigcap_{i=1}^{4} B_i\right) = 0$$

Using general addition rule:

$$P(A^c) = 4\cdot\left(\frac{3}{4}\right)^6 - {}^4C_2\left(\frac{1}{2}\right)^6 + 4\cdot\left(\frac{1}{4}\right)^6$$

$$= \frac{317}{512} \simeq 0.62$$

$$P(A) = \frac{195}{512} \simeq 0.38$$

Example: Three players A, B, C take turns in throwing a dice in order A, B, C, . . . A, B, C . . , ... What is the prob that A is the second player to get a six for the first time?

Sol^n: The player A gets a chance to throw the dice on $(3r+1)^{th}$ trial, $r = 0, 1, 2 \ldots$

So in order that he is second to

throw a six, it can be on any of $(3r+1)^{th}$ trial, $r = 1, 2, \ldots$. On $(r+1)$ trials that A gets to throw, r are not six and the last is 6. So the prob is

$$\left(\frac{5}{6}\right)^r \cdot \frac{1}{6}.$$

B may get a six on r trials w.p. $1-\left(\frac{5}{6}\right)^r$

C may not get any six on r trials w.p. $\left(5/6\right)^r$

So prob that A throws a six after B but before C on $(3r+1)^{th}$ trial is

$$\left(\frac{5}{6}\right)^{2r}\left\{1-\left(\frac{5}{6}\right)^{r}\right\}\cdot\frac{1}{6}$$

Similar expression for C first & B last

So P(A is second to throw a six)

$$= 2\sum_{r=1}^{\infty}\left(\frac{5}{6}\right)^{2r}\left\{1-\left(\frac{5}{6}\right)^{r}\right\}\cdot\frac{1}{6}$$

$$= \frac{300}{1001} \simeq 0.2997.$$

(ii) What is the prob that A is last to get a fix for the first time (first)