Random Experiment

A random experiment may be defined as an experiment with more than one out come, which may be repeated any number of times under more or less similar conditions and the outcomes of which varies irregularly from repetition to repetation

So A random experiment is an experiment in which:

(a) All out come of an experiment is known in advance (and is denoted by Ω)

(b) Any performance of the experiment results in an ~~experiment~~ out come that is not known in advance

(c) The experiment can be repeated under identical Conditions

σ- field of events or set (S)

Any ~~set~~ Sub set of the set Ω (set of all possible out come of the experiment) is called an event.

Any set S of the events is called a σ-field of events if it has the following properties

1) The set Ω and the null set ϕ are in S
ie $\Omega \in S$ and $\phi \in S$

2) If a finite or countable sequence of events $A_1 A_2 \cdots$ are in ~~B~~ S then their union and their intersection are in S, i.e if $A_i \in S$. Then $A_1 \cap A_2 \cap \cdots \in S$ and $A_1 \cup A_2 \cup \cdots \in S$.

3) If A is an event in S its complement A^c will also be in S ie $A \in S \Rightarrow A^c \in S$.

Eg:

Let $\Omega = \{1, 2, 3\}$

Then $S = \{\phi, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$.

is a σ field

Sample space

The sample space of a statistical experiment is pair (Ω, S), where

a) Ω is the set of all posible out come of the experiment

b) S is a σ field of subset of Ω

Different type of classical probabilities

The three type of classical probabilities are

1) Malthematical Probability
2) statistical regularity or Von Mises's Probability
3) Axiomatic Probability or Probability measure.

Mathematical Probability

If Consistent with the given condition. let there are n exhaustive, mutully exclusive and equaly likely cases in Ω and m of them are favourable to an event A then the mathematical probability of A is defined as m/n.

Von Mises's Statistical or Emperical probability

If trials be repeated a greater number of times under, essentialy the same conditions, the limit of the ratio of number of times that an event happends to the total number of trials as the number of trials increases indefinitely is called the probability of happening of that event.

Probability Measure

Function Two sets D and R and a rule which assign a unique element of R to every element of D defines a function (say f)

Here D is called the domain and R is called the Range of the function

Set function: If the domain of a function is a set of sets, the ~~fun~~ function is called a set function

Real valued functions: If the range set R of a function is a set of real numbers, the function is called real valued set function

Additive and Sigma additive set function

If A and B are any two disjoint sets in a domain of real valued function f and if

$f(A \cup B) = f(A) + f(B)$.

the set function is said to be additive

If $A_1, A_2 \ldots$ is an enumerable sequence of disjoint sets on the domain of set function f and if

$f(A_1 \cup A_2 \cup \ldots) = f(A_1) + f(A_2) + \cdots$

then the set function f is said to be totaly additive or sigma additive

Measure

A real valued totaly additive set function is called a measure.

Probability Measure

Let (Ω, S) be a sample space. A set function P defined ~~over~~ on S is called a probability measure if it satisfies the following conditions

1) $P(A) \geq 0$ for all $A \in S$

2) $P(\Omega) = 1$.

3) Let $\{A_j\}$, $A_j \in S$, $j = 1, 2 \ldots$ be disjoint sequences of sets that is, $A_i \cap A_j = \phi$ for all $i \neq j$, where ϕ is the null set. Then

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

In other words

A real valued non negative totaly additive set function P defined on the sample space (Ω, S) is called a probability measure if it satisfies the condition $P(\Omega) = 1$, and is denoted by. (Ω, S, P).

The triple (Ω, S, P) is called Probability space.

<u>Definition</u> Let $A \in S$. We say that odds for A are a to b if $P(A) = \frac{a}{a+b}$ and then the odds against A are b to a

<u>Theorem 1</u>

P is monotone and sbtractive: that is if $A, B \in S$ and $A \subseteq B$, Then $P(A) \leq P(B)$ and $P(B-A) = P(B) - P(A)$ where. $B - A = B \cap A^c$, A^c being the complement of the event A.

<u>Proof</u>

For ~~any~~ the events A and B we have

$$B = (A \cap B) \cup (B \cap A^c) \quad [\text{or } B = (A \cap B) \cup (B - A)]$$
$$= A + (B \cap A^c) \quad \text{Since } A \subseteq B,\ A \cap B = A.$$

~~Also~~ Also $A \cap (B \cap A^c) = \phi$. So by axiom 3 we have

$$P(B) = P(A) + P(B \cap A^c) = P(A) + P(B-A) \quad \text{—— (1)}$$
$$\Rightarrow P(B-A) = P(B) - P(A).$$

Now by axiom (1) $P(A) \geq 0$ and $P(B-A) \geq 0$

So from (1) we have $P(B) \geq P(A)$

<u>Corollary</u> For all $A \in S$, $0 \leq P(A) \leq 1$

By Axiom (1) for any $A \in S$ we have $P(A) \geq 0$ — (1)

By theorem (1) if $A \subseteq B$ then $P(A) \leq P(B)$

Now take $B = \Omega$. Then we have $P(A) \leq P(\Omega) = 1$ — (2)
by axiom (2).

Combaining (1) and (2) we get $0 \leq P(A) \leq 1$.

<u>Notations</u>

$$A \cup B = A + B.$$

$$A \cap B = AB$$

$$A \cap B^c = A - B$$

$$A_1 \cup A_2 \cup \cdots \cup A_n = A_1 + A_2 + \cdots + A_n = \sum_{i=1}^{n} A_i$$

ie $$\bigcup_{i=1}^{n} A_i = \sum_{i=1}^{n} A_i$$

$$A_1 \cap A_2 \cap \cdots \cap A_n = A_1 A_2 \cdots A_n = \prod_{i=1}^{n} A_i$$

ie $$\bigcap_{i=1}^{n} A_i = \prod_{i=1}^{n} A_i$$

<u>Theorem 2</u> (The Addition Rule)

If $A, B \in S$ then

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

or

$$P(A+B) = P(A) + P(B) - P(AB).$$

Proof

For any two events A and B we have

$B = (A \cap B) \cup (B \cap A^c)$ ——— (1)

$A \cup B = A \cup (B \cap A^c)$ ——— (2)

Now $(A \cap B) \cap (B \cap A^c) = \phi$ and $A \cap (B \cap A^c) = \phi$.

So from (2) using axiom 3 we can write

$P(A \cup B) = P[A \cup (B \cap A^c)] = P(A) + P(B \cap A^c)$ ——— (3)

and from (1) we have

$P(B) = P(A \cap B) + P(B \cap A^c)$

or $P(B \cap A^c) = P(B) - P(A \cap B)$ ——— (4)

Subtituting (4) in (3) we get

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ——— (5)

Corollary (1): P is subaddetive, that is if $A, B \in S$ then $P(A \cup B) \leq P(A) + P(B)$

Corollary (2) If $B = A^c$ then A and B are disjoint and hence

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

[For this we have $A \cup A^c = \Omega$ or $P(A \cup A^c) = P(\Omega) = 1$ and $P(A \cap A^c) = P(\phi) = 0$ Subtitituting this in (5) of theorem we have the Corollary]

Note

Corollary (1) Can be extented to any arbitory number of A_j's ie $P(\bigcup_j A_j) \leq \sum_j P(A_j)$

Theorem 3 (Principle of Inclusion - Exclusion)

Let $A_1, A_2 \ldots A_n \in \mathcal{S}$ Then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i \\ i<j}} \sum_{j} P(A_i A_j)$$

$$+ \sum_{\substack{i \\ i<j<k}} \sum_{j} \sum_{k} P(A_i A_j A_k)$$

$$+ \cdots + (-1)^{n} P\left(\bigcap_{i=1}^{n} A_i\right) \quad \text{—— (1)}$$

Proof

The proof is by mathematical induction

Lets consider the case $n=2$. Then ~~Now~~ by theorem 2 we have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2) \quad \text{—— (2)}$$

Now let us assume that the result holds for $n=m$ ie we have

$$P\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} P(A_i) - \sum_{i}\sum_{j} P(A_i A_j) - \sum_{\substack{i \\ i<j<k}}\sum_{j}\sum_{k} P(A_i A_j A_k)$$

$$+ \cdots + (-1)^{m} P(A_1 A_2 \cdots A_m) \quad \text{—— (3)}$$

Now let $W = \bigcup_{i=1}^{m} A_i$ then using (2) we have

$$P(W \cup A_{m+1}) = P(W) + P(A_{m+1}) - P(W A_{m+1})$$

substituting for W we get

$$P\left(\bigcup_{i=1}^{m} A_i \cup A_{m+1}\right) = P\left(\bigcup_{i=1}^{m} A_i\right) + P(A_{m+1}) - P\left[\left(\bigcup_{i=1}^{m} A_i\right) \cap A_{m+1}\right] \quad \text{—— (4)}$$

Combaining (3) and (4) we get

$$P\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} P(A_i) - \underset{i \le j \le m}{\sum_i \sum_j} P(A_i A_j) + \cdots + (-1)^{m} P\left(\bigcap_{i=1}^{m} A_i\right)$$

$$+ P(A_{m+1}) - \Big[\sum_{i=1}^{m} P(A_i A_{m+1}) - \cancel{\sum\sum_{i<j<k<m} P(A_i A_j A)}$$

$$- \sum_{i<j<k<m} P(A_i A_j A_{m+1}) + \cdots$$

$$+ (-1)^{m+1} P(A_1 A_2 \cdots A_m) \Big]$$

$$= \sum_{i=1}^{m+1} P(A_i) - \underset{i<j<m+1}{\sum_i \sum_j} P(A_i A_j) + \cdots$$

$$+ (-1)^{m+1} P\left(\bigcap_{i=1}^{m+1} A_i\right)$$

So the result holds for $n = (m+1)$. So by induction the result holds for any positive intiger $n \geq 2$.

Theorem 4 (Bonferronis Inequality)

Given $n (\geq 1)$ events $A_1\ A_2 \cdots A_n$ satisfies

$$\sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) \le P\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i) \quad \text{---} (1)$$

Proof

The proof is by induction

Consider the case $n=2$. Then by addition Theorem-2 we have

$$P(A_1) + P(A_2) - P(A_1 A_2) = P(A_1 \cup A_2) \quad \text{---} (2)$$

So the L HS of the inequality holds.

for RHS

we have again $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2)$

and $P(A_1 A_2) \geq 0$ so $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$

or $P(A_1 \cup A_2) \leq \sum_{i=1}^{2} P(A_i)$

So the inequality holds (1) for $n=2$.

For $n \geq 3$ ~~the~~ using the theorem (3) the RHS ~~R~~ inequality holds. So it is need only to prove the LHS inequality only.

Now ~~for n=~~ assume that inequality is true for $n=m$ ie we have

$$P\left(\bigcup_{i=1}^{m} A_i\right) \geq \sum_{i=1}^{m} P(A_i) - \sum_{i \neq j} P(A_i A_j) \quad \text{———} (3)$$

Let $W = \bigcup_{i=1}^{m} A_i$ Then

$$P\left(\bigcup_{i=1}^{m+1} A_i\right) = P(W \cup A_{m+1}) = P(W) + P(A_{m+1}) - P(W A_{m+1})$$

$$= P\left[\left(\bigcup_{i=1}^{m} A_i\right) \cup A_{m+1}\right]$$

$$= P\left(\bigcup_{i=1}^{m} A_i\right) + P(A_{m+1}) - P\left[\left(\bigcup_{i=1}^{m} A_i\right) \cap A_{m+1}\right]$$

[using (2)]

$$\geq \sum_{i=1}^{m} P(A_i) - \sum_{i \neq j} P(A_i A_j) + P(A_{m+1})$$

$$- \sum_{i=1}^{n} P(A_i \cap A_{m+1}) \quad [\text{using } (3)]$$

$$\geq \sum_{i=1}^{m+1} P(A_i) - \sum_{i<j}^{m+1} P(A_i A_j)$$

So the result holds for ~~all~~ $n = m+1$. Hence the result holds for all positive integer $n > 1$.

Theorem 5 (Boole's Inequality) For any two events A and B

$$P(A\cap B) \geq 1 - P(A^c) - P(B^c). \quad \text{—— (1)}$$

Proof

We have $(A\cap B) \cup (A\cap B)^c = \Omega$ and $(A\cap B)\cap(A\cap B)^c = \phi$

So by Axioms 2 and 3 we have

$$P[(A\cap B) \cup (A\cap B)^c] = P(\Omega)$$

ie $P(A\cap B) + P(A\cap B)^c = 1.$

$$P(A\cap B) = 1 - P(A\cap B)^c \quad \text{—— (2)}$$

Now by De Morgan's law we have

$$(A\cap B)^c = (A^c \cup B^c)$$

So $P(A\cap B)^c = P(A^c) + P(B^c) - P(A^cB^c)$ —— (3)

Substituting (3) in (2) we get

$$P(A\cap B) = 1 - [P(A^c) + P(B^c) - P(A^cB^c)]$$

$$= 1 - P(A^c) - P(B^c) + P(A^cB^c)$$

So

$P(A\cap B) \geq 1 - P(A^c) - P(B^c)$. Since $P(A^cB^c) \geq 0$

Corollary 1 Let $\{A_j\}$, $j = 1, 2, \ldots$ be a countable sequence of events then $P(\cap A_j) \geq 1 - \Sigma P(A_j^c)$

Proof

Take $B = \bigcap_{j=2}^{\infty} A_j$ and $A = A_1$ in Booles inequality

Corollary 2 (Implication Rule) : If $A, B, C \in S$ and A and B imply C, then $P(C^c) \leq P(A^c) + P(B^c)$

Limit Superior and limit inferior

Let $\{A_n\}$ be a sequence of sets. The set of all point $\omega \in \Omega$ that belong to A_n for infinitely many value of n is known as the limit superior of the sequence and is denoted by

$$\limsup_{n \to \infty} A_n \text{ or } \overline{\lim_{n \to \infty}} A_n$$

The set of all points that belong to A_n for all but a finite number of values of n is known as the limit inferior of the sequence $\{A_n\}$ and is denoted by

$$\liminf_{n \to \infty} A_n \text{ or } \underline{\lim_{n \to \infty}} A_n$$

Limit of the Sequence

If $\underline{\lim_{n \to \infty}} A_n = \overline{\lim_{n \to \infty}} A_n$, we say that limit exist and write $\lim_{n \to \infty} A_n$ for the common set and call it limit set.

Note

we have $\underline{\lim_{n \to \infty}} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \subseteq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \overline{\lim_{n \to \infty}} A_n$

Nondecreasing and Non increasing sequence.

If the sequence $\{A_n\}$ is such that $A_n \subseteq A_{n+1}$ for $n = 1, 2, \ldots$ it is called non decreasing sequence and is denoted by $A_n \uparrow$

Simillarly if the sequence is such that $A_n \supseteq A_{n+1}$ for $n = 1, 2, \ldots$ it is called non increasing sequence and is denoted by $A_n \downarrow$.

Note

If $\{A_n\}$ is $A_n \uparrow$ or $A_n \downarrow$, then the limit exists and is given by

$$\lim_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} A_n \quad \text{if } A_n \downarrow$$

and

$$\lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n \quad \text{if } A_n \uparrow$$

Theorem 6

Let $\{A_n\}$ be a non decreasing sequence of events in S, ie $A_n \in S$, for $n = 1, 2, \ldots$ and $A_n \supseteq A_{n-1}$ for $n = 2, 3, \ldots$

Then.

$$\lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right) \quad \text{—(1)}$$

Proof

Let $A = \bigcup_{j=1}^{\infty} A_j$

Then

$$A = A_n + \sum_{j=n}^{\infty} (A_{j+1} - A_j)$$

Note

$[A_{j+1} - A_j = A_{j+1} \cap A_j^c]$

By countable additivity we have

$$P(A) = P(A_n) + \sum_{j=n}^{\infty} P(A_{j+1} - A_j)$$

So as $n \to \infty$ we have

$$P(A) = \lim_{n\to\infty} P(A_n) + \lim_{n\to\infty} \sum P(A_{J+1} - A_J)$$

Now since $\sum_{J=1}^{\infty} P(A_{J+1} - A_j) \leq 1$ and each summand is non-negative, the second term on RHS of (2) tends to zero. So we have

$$P(A) = \lim_{n\to\infty} P(A_n)$$

But $A = \bigcup_{J=1}^{\infty} P(A_j)$

So $$\lim_{n\to\infty} P(A_n) = P\left(\bigcup_{J=1}^{\infty} A_j\right).$$

<u>Corrollary</u> Let $\{A_n\}$ be a non increasing sequence of events in S. Then $\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P\left(\bigcap_{J=1}^{\infty} A_J\right)$ — (1)

<u>Proof</u>

Consider the non decreasing sequence of events $\{A_c^c\}$. Then $\lim_{n\to\infty} A_n^C = \left(\bigcup_{J=1}^{\infty} A_n\right)^C = A^C$

So using theorem 6 we have

$$\lim_{n\to\infty} P(A_n^c) = P(\lim_{n\to\infty} A_n^c) = P\left(\bigcup_{J=1}^{\infty} A_n^c\right) = P(A^C)$$

which is equivalent to

$$\lim_{n\to\infty} (1 - P(A_n)) = 1 - P(A) \Rightarrow \lim_{n\to\infty} P(A_n) = P(A)$$

Now by Demorgans law $A^c = \left(\bigcup_{J=1}^{\infty} A_n^c\right) = A = \left(\bigcap_{J=1}^{\infty} A_J\right)$

Hence the Corollary.

Conditional Probability

Let (Ω, S, P) be a probability space, and let $H \in S$ with $P(H) > 0$. For any arbitary $A \in S$, the conditional probability of A given H is defined as

$$P(A|H) = \frac{P(A \cap H)}{P(H)}, \quad P(H) > 0$$

Conditional probability remains undefined when $P(H) = 0$

Theorem - 1

Let (Ω, S, P) be a probability space and let $H \in S$ with $P(H) > 0$. Then (Ω, S, P_H), where $P_H(A) = P\{A|H\}$ for all $A \in S$, is a probability space.

Proof

$P_H(A) = P\{A|H\} \geq 0$ for all $A \in S$ as $P(A \cap H) \geq 0$ and $P(H) > 0$

~~Also~~ $P(\Omega|H) = \frac{P(\Omega \cap H)}{P(H)} = \frac{P(H)}{P(H)} = 1$ as $H \subseteq \Omega$

If $A_1, A_2, \dots$ is a disjoint sequence set of S, then

$$P_H \left(\bigcup_{i=1}^{\infty} A_i \right) = P\left\{ \bigcup_{i=1}^{\infty} A_i \Big| H \right\} = \frac{P\{(\bigcup_{i=1}^{\infty} A_i) \cap H\}}{P(H)}$$

$$= \frac{P\left[\bigcup_{i=1}^{\infty} (A_i \cap H)\right]}{P(H)} = \sum_{i=1}^{\infty} \frac{P(A_i \cap H)}{P(H)}$$

$$= \sum_{i=1}^{\infty} P_H(A_i)$$

So $P_H(A)$ satisfies all the three condition of Probability measure. Hence (Ω, S, P_H) is a probability measure.

Result

Let A and B be two events in S with $P(A) > 0$ and $P(B) > 0$. Then by definition of conditional probability we have

$$P(A \cap B) = P(B)\, P(A \mid B) \quad \text{and} \quad P(A \cap B) = P(A)\, P(B \mid A)$$

Theorem 8 (Multiplication Rule)

Let (Ω, S, P) be a probability space and $A_1, A_2, \dots, A_n \in S$ with $P(\bigcap_{j=1}^{n-1} A_j) > 0$, Then

$$P(\bigcap_{j=1}^{n} A_j) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 A_2) \dots P\{A_n \mid A_1 A_2 \dots A_{n-1}\} \quad \text{--- (1)}$$

Proof

The proof is by induction method

Consider the two event A_1 and A_2. $(n=2)$ Then by definition of conditional probability we have

$$P(A_2 \mid A_1) = \frac{P(A_1 A_2)}{P(A_1)} \quad \text{--- (2)}$$

So we have $P(A_1 A_2) = P(A_1)\, P(A_2 \mid A_1)$ — (3)

So the result holds for $n = 2$. Now Assume that the result holds for $n = m$ and $W = \bigcap_{j=1}^{m} A_j$. Then by (1) we have

$$P(W) = P(\bigcap_{j=1}^{m} A_j) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 A_2) \dots P\{A_m \mid A_1 A_2 \dots A_{m-1}\} \quad \text{--- (4)}$$

Now using (3) we have

$$P(W \cap A_{m+1}) = P(W)\,P(A_{m+1} | W)$$

Now Substituting for W we get

$$P\left[\left(\bigcap_{j=1}^{m} A_j\right) \cap A_{m+1}\right] = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2) \cdots P(A_m|A_1A_2\cdots A_{m-1})\,P(A_{m+1}|A_1A_2\cdots A_m)$$

ie $$P\left(\bigcap_{j=1}^{m+1} A_j\right) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2) \cdots P(A_{m+1}|A_1A_2\cdots A_m)$$

So the result holds for $n = m+1$. Hence the result hold for any $n > 1$ by mathematical Induction

<u>Theorem 9</u> (Bayes Rule)

Let $\{H_n\}$ be a sequence of disjoint events such that $P\{H_n\} > 0$, $n = 1, 2, \ldots$ and $\sum_{n=1}^{\infty} H_n = \Omega$. Let $B \in S$ with $P(B) > 0$. Then

$$P\{H_j | B\} = \frac{P(H_j)\,P(B|H_j)}{\sum_{j=1}^{\infty} P(H_j)\,P(B|H_j)} \qquad j = 1, 2, \ldots \qquad (1)$$

<u>Proof</u>

By the definition of conditional probability we have $P(B \cap H_j) = P(B)\,P(H_j|B) = P(H_j)\,P(B|H_j)$.

So $$P(H_j|B) = \frac{P(H_j)\,P(B|H_j)}{P(B)} \quad \text{———} \quad (2)$$

Since H_j . $j = 1, 2 \dots$ are disjoint events in (Ω, S, P)

we can have $B = H_1 B \cup H_2 B \cup \dots$

$$P(B) = P[H_1 B \cup H_2 B \cup \dots] = \sum_{i=1}^{\infty} P(H_i \cap B)$$

$$= \sum_{i=1}^{\infty} P(H_i) P(B|H_i) \quad \text{—— (3)}$$

Subtituting (3) in (2) we get

$$P(H_j | B) = \frac{P(H_j) P(B|H_j)}{\sum_{i=1}^{\infty} P(H_j) P(B|H_j)}$$

Hence the theorem

<u>Independence of Events</u>

Let (Ω, S, P) and let $A, B \in S$. Then A and B are said to be independent if and only if

$$P(AB) = P(A) P(B).$$

<u>Theorem - 10</u>

If A and B are independent events then

$$P(A|B) = P(A) \qquad \text{if } P(B) > 0$$

and

$$P(B|A) = P(B) \qquad \text{if } P(A) > 0.$$

<u>Proof</u>

Given A and B are independent so

$P(AB) = P(A) P(B)$.

So $P(A|B) = \dfrac{P(AB)}{P(B)} = \dfrac{P(A)P(B)}{P(B)} = P(A)$

and $P(B|A) = \dfrac{P(AB)}{P(A)} = \dfrac{P(A)P(B)}{P(A)} = P(B)$

Theorem-11

If A and B are independent, so are A and B^c, A^c and B and A^c and B^c.

Proof

Give A and B are independent so we have
$P(AB) = P(A)P(B)$, $P(A|B) = P(A)$ and $P(B|A) = P(B)$

Now
$$P(AB^c) = P(A)P(B'|A)$$
$$= P(A)[1 - P(B|A)]$$
$$= P(A)[1 - P(B)]$$
$$= P(A)P(B^c).$$

So A and B^c are independent. Simillerly we can prove that A^c and B are independent

Now $P(A^cB^c) = P(A^c)P(B^c|A^c)$
$$= P(A^c)[1 - P(B|A^c)]$$
$$= P(A^c)[1 - P(B)]$$ Since B and A^c are independent
$$= P(A^c)P(B^c)$$

Hence A^c and B^c are independent.

Pairwise independence. Let $\mathcal{U}$ be a family of events from S. We say that the events in $\mathcal{U}$ are pairwise independent if and only if for every pair of distinct events, $A, B \in \mathcal{U}$

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

Mutual or Complete independence

A family of events $\mathcal{U}$ is said to be a mutully or Completely independent family if and only if for every finite subcollection $\{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ of $\mathcal{U}$ the following relation holds.

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_j})$$
$$= \prod_{j=1}^{k} P(A_{i_j})$$

Note

This has $2^n - n - 1$ relations to be hold for independence. as follows

$$P(A_i A_j) = P(A_i) P(A_j) \quad i \neq j,\ i, j = 1, 2, \ldots, n$$
$$P(A_i A_j A_k) = P(A_i) P(A_j) P(A_k) \quad i \neq j \neq k,\ i, j, k = 1, 2, \cdots n$$
$$\vdots$$
$$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2) \cdots P(A_n).$$

Counter Example

Take four marbles identical in size, one the first write $A_1 A_2 A_3$ and on each of the other three write, A_1, A_2, A_3, respectively. Put 4 marbls in an Urn and draw one at random. Let E_i denote the event that the symbol A_i appear on the drawn marbles. Then

$$P(E_1) = P(E_2) = P(E_3) = 1/2$$
$$P(E_1 E_2) = P(E_1 E_3) = P(E_2 E_3) = 1/4$$

and $P(E_1 E_2 E_3) = 1/4 \neq P(E_1) P(E_2) P(E_3)$. So $E_1 E_2 E_3$ are not independent even though they are pairwise independent.

Borel 0-1 Crieterian

Let $\{A_n\}$ be a sequence of independent events Then $P(\overline{\lim} A_n)$ is zero or one and $P(\underline{\lim} A_n)$ is either zero or one. So if both $P(\underline{\lim} A_n) = P(\overline{\lim} A_n) = P(A)$ exist Then $P(A)$ is either zero or one. The following theorem gives the precise criterian for $P(\overline{\lim} A_n)$ to be either zero or one and is called Borel a.s. Criterion.

Theorem

Let $\{A_n\}$ be a sequence of ~~independent~~ event and

(1) If $\Sigma P(A_n) < \infty$, $P(\overline{\lim} A_n) = 0$

(2) If $\Sigma P(A_n) = \infty$ and A_n's are independent,

$$P(\overline{\lim} A_n) = 1.$$

Proof

The first part of the theorem is true without the assumption of independence of the A_k's. The first part is also called Borel - Cantelli Lemma.

1) Since $P(\cup A_k) \leq \Sigma P(A_k)$, whether A_k's are independent or not,

$$P\left(\bigcup_{r}^{s} A_k\right) \leq \sum_{r}^{s} A_k \to 0$$

as $r, s \to \infty$ if $\Sigma P(A_n) < \infty$. But this implies that

$$\lim_{r} \lim_{s} P\left(\bigcup_{r}^{s} A_k\right) = P(\overline{\lim} A_k) = 0$$

2) To prove the second part, we have

$$P(\overline{\lim} A_n) = \lim_{r} \lim_{s} \left[1 - P\left(\bigcap_{r}^{s} A_k^c\right)\right]$$

$$= \lim_{r} \lim_{s} \left[1 - \prod_{k=r}^{s} (1-P(A_k))\right] \quad \text{--- (1)}$$

Since $\{A_k^c\}$ are independent and

$$P\left(\bigcap_{k} A_k^c\right) = \prod_{k} (1-P(A_k))$$

But for $0 \le P(A_k) \le 1$, $1-P(A_k) \le \exp(-P(A_k))$ so

$$\prod_{k} (1-P(A_k)) \le \exp\left(-\sum_{k} P(A_k)\right)$$

and therefore, $1 - \exp\left(-\sum_{k} P(A_k)\right) \le 1 - \prod_{k} (1-P(A_k))$ — (2)

If $\sum P(A_n) = \infty$, then $\sum_{k=n}^{\infty} P(A_k)$ is infinite for all n.

Then $\lim_{r} \lim_{s} \sum_{r}^{s} P(A_k) = \infty$

$$\text{ie} \quad \lim_{r} \lim_{s} \exp\left\{-\sum_{k=r}^{s} P(A_k)\right\} = 0 \quad \text{--- (3)}$$

using (1), (2) and (3) we have

$$P(\overline{\lim} A_n) = 1.$$

Hence the theorem.