PROBABILITY

Sample space and events:

The set of all possible outcomes of a particular experiment is called the sample space. Denoted by S.

The subset of S is known as an event.

E.g S = Rolling a dice
Possible outcomes = $\{1, 2, 3, 4, 5, 6\}$

E = No. is even
Outcome = $\{2, 4, 6\} \subseteq S$

If there are two events E & F, $E \cup F$ (E union F) consists of all outcomes that are either in E or F or both. $E \cap F$, (EF) (E intersection F) consists of common outcomes in E and F. ϕ refers to no outcomes for that particular event. E^c refers to complement of E consists of all ~~events~~ outcomes in S but not in E. Similarly, $E \subset F$: outcomes in E present in F. $E \subset F \Rightarrow E \supset F \Rightarrow E = F$.

Axioms of probability:

Axiom 1: $P(E) = \dfrac{\text{No. of outcomes in } E}{\text{No. of outcomes in } S}$

Axiom 2: $0 \leq P(E) \leq 1$

Axiom 3: $P(S) = 1$

Axiom 4: For any sequence of mutually exclusive events $E_1, E_2, E_3, E_4, \ldots$ (Mutually exclusive = $E_i \cap E_j = \phi$ when $i \neq j$)

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) \qquad n = 1, 2, 3, \cdots \infty$$

Proposition 1: $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

odds of an event A $= \dfrac{P(A)}{P(A^c)} = \dfrac{P(A)}{1 - P(A)}$

<u>Conditional Probability</u>: Very useful when partial information is given concerning the result of the experiment is available.

If there are two events E and F, so probability of E given that F has occured

$$= P(E|F) = \frac{P(E \cap F)}{P(F)} \quad P(F) > 0 \quad -① $$

Suppose if the experiment is performed n times (n is large).

$\because$ P(F) is long-run proportion of experiment in which F occurs

F = n P(F) similarly n P(EF).

Hence,

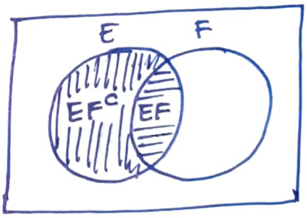
$$\left.\begin{array}{l} nP(F) \xrightarrow{\text{outcome}} \text{in } F \\ n(P(EF)) \xrightarrow{\text{outcome}} \text{in } F \end{array}\right\} \text{approx.}$$

So, now $p'(E|F) = \frac{\cancel{n} P(EF)}{\cancel{n} P(F)} = \frac{n P(EF)}{P(F)} = P(E|F)$

From eq. 1 $\quad P(EF) = P(E|F) \cdot P(F)$

<u>Bayes' Formula</u>:

Let E & F be two events,



Any event E can be expressed as

$E = EF \cup EF^c$

$\because$ EF and EF^c are mutually exclusive,

$P(E) = P(EF) + P(EF^c)$

$\quad = P(F) \cdot P(E|F) + P(F^c) P(E|F^c)$

$\quad = P(F) \cdot P(E|F) + (1 - P(F)) P(E|F^c)$

Theoritically, P(E) = Weighted average of conditional probability of E given F has occured and E given F hasn't occured.

Total Probability. If event A is to be found given that events $B_1, B_2, B_3, \cdots B_n$ have occurred.

$$P(A) = P((A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cdots \cup (A \cap B_n))$$

$$= \sum_{k=1}^{n} P(A \cap B_k) = \sum_{K=1}^{n} P(B_K)\, P(A \mid B_K)$$

Bayes' Rule (Extended)

If $B_1, B_2, B_3, \cdots B_n$ form a partition of sample space, then for event A with $P(A) > 0$,

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)\, P(A \mid B_i)}{P(B_1)\, P(A \mid B_1) + P(B_2)\, P(A \mid B_2) + \cdots + P(B_n) P(A \mid B_n)}$$

Independent events :

There are case when one event E doesn't affect the event F. In these situations, Events E and F are called independent events.

Now,

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

$$P(E \cap F) = P(E \mid F) \cdot P(F)$$

$\because$ E & F are independent

$$P(E \mid F) = P(E)$$

$$\therefore P(E \cap F) = P(E) \cdot P(F)$$

Proposition: If E & F are independent, so is E and F^c.

Proof: Assume E and F are independent.

$$P(E) = P(EF \cup EF^c)$$

$$= P(EF) + P(EF^c)$$

$$= P(F)\, P(E \mid F) + P(F^c)\, P(E \mid F^c)$$

$$= P(F) \cdot P(E) + \underbrace{P(F^c)\, P(E \mid F^c)}_{P(EF^c)}$$

Equivalently : $P(EF^c) = P(E) - P(E) \cdot P(F)$

$$= P(E)\,[1 - P(F)] = P(E) \cdot P(F^c)$$

Random Variables: A variable whose value is unknown or a function that assigns a value to each of an experiment's outcomes.

E.g E = Rolling two dice

X = Random variable # Sum of two fair dice.

$X = x$	2	3	4	5	6	7	8	9	10	11	12
$P(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Observing these values,

$$\sum_{i=1}^{11} P(X = x_i) = P(S) = 1.$$

Cumulative distribution function: or F of the random variable X is defined for any $x \in \mathbb{R}$ by

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

Denotion of cumulative distribution function: $X \sim F$

Discrete random variable: A variable that can take any whole number values as outcomes of a random experiment.

$$\Omega_X = \underbrace{\{x_1, x_2, x_3, \cdots\}}_{\text{Countable set.}} \qquad X : \Omega \rightarrow \mathbb{R}.$$

Probability mass function: For a discrete random variable X, we define PMF $p(a)$ of X by,

$$p(a) = P\{X = a\}$$

$p(a) > 0$ for atmost 'a' countable values

i.e X must assume one of $x_1, x_2, \ldots$ then

$p(x_i) > 0$ and $p(x) = 0$

$i = 1, 2, 3, \ldots$ $\quad x \rightarrow$ all other values.

$\because X$ must have one of values x_i

$$\therefore \sum_{i=1}^{\infty} p(x_i) = 1$$

Jointly Distribution random variables : When we deal with 2 or more random variables.

PMF of X and Y = $F(x,y) = P\{X \le x, Y \le y\}$

In case of discrete random variables,

where $X = x_1, x_2, x_3, \ldots$

$Y = y_1, y_2, y_3, \ldots$

Joint PMF $p(x_i, y_i) = P\{X = x_i, Y = y_i\}$

* Individual PMFs obtained by Joint PMF.

Proof: $\because$ Y must take some y_i, follows that event $\{X = x_i\}$ can be written as union over all j of the mutually exclusive events

$X = x_i, Y = y_j$

$$\{X = x_i\} = \bigcup_j \{X = x_i, Y = y_j\}$$

$$P(\{X = x_i\}) = P\left(\bigcup_j \{X = x_i, Y = y_j\}\right)$$

$$= \sum_j P\{X = x_i, Y = y_j\}$$

$$= \sum_j p(x_i, y_i)$$

Same can be applied for obtaining $P\{Y = y_j\} = \sum_j p(x_i, y_i)$

However, the converse is not true.

Independent random variables

Random variable X and Y are said to be independent if for any two sets of real numbers A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}$$

Theoretically, if $E_A = \{X \in A\}$ and E_B $\{Y \in B\}$ are independent then X and Y are independent.

In terms of joint PMF of X and Y,

$F(a, b) = F_X(a) F_Y(b)$ for all a, b.

$= p(x,y) = p_X(x) \cdot p_Y(y)$

Now

$$P\{X \in A, Y \in B\} = \sum_{x \in A} \sum_{y \in B} p(x,y)$$

$$= \sum_{y \in B} \sum_{x \in A} p_X(x) \cdot p_Y(y)$$

$$= \sum_{x \in A} p_X(x) \cdot \sum_{y \in B} p_Y(y)$$

$$= P(Y \in B) \cdot P(X \in A).$$

Conditional PMF: The relationship between two random variables can be clarified by consideration of conditional distribution of one given the value of other

Similar to conditional probability, if X, Y are discrete R.V, conditional PMF of X given $Y = y$,

$$p_{X|Y}(x|y) = P(X = x \mid Y = y)$$

$$= \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x,y)}{p_Y(y)} \qquad p_Y(y) > 0.$$

Expectation: If X is a discrete random variable taking on possible values $x_1, x_2, \ldots$ then expectation or expected value of X denoted by $E(X)$

$$E(X) = \sum_i x_i \, P\{X = x_i\}$$

$$= \sum_i x_i \, p_X(x_i)$$

Expected value of a function of X ($g(X)$): Suppose X is given and we have to find the expected value of a function of X ($g(X)$). To find this, $\because$ $g(X)$ itself is random variable $\longrightarrow$ PMF of $g(X)$ computable with the help of X, then compute $E(g(X))$

$$= \sum_{i=1}^{n} x_i \, P\{Y = g(X)\} = \sum_x g(x) \, p(x)$$

E.g: Suppose X has PMF

$p(0) = 0.2$, $p(1) = .5$, $p(2) = .3$

calculate $E(X^2)$

Ans. If $Y = X^2$

$\because$ Y is random variable, it can take one of values $0^2, 1^2, 2^2$ with resp. probabilities.

$p_Y(0) = p(Y = 0^2) = 0.2$

$p_Y(1) = p(Y = 1^2) = 0.5$

$p_Y(2) = p(Y = 2^2) = 0.3$

$\therefore E(X^2) = \sum x_i P(Y = x^2)$

$= 0(0.2) + 1(0.5) + 4(0.3)$

$= 0.5 + 1.2 = 1.7.$

<u>Expected value of sums of random variables</u>:

If X and Y are two random variables and g is a function of two variables, then

$$E(g(X,Y)) = \sum_y \sum_x g(x,y) f(x,y)$$

In general,

$$E[X_1 + X_2 + X_3 + \dots X_n] = E[X_1] + E[X_2] + \dots + E[X_n]$$

<u>Variance</u>: If X is a random variable, then the Var of X is denoted by (Variation refers to the deviation of a random variable from its mean)

$$\text{Var}(X) = E((X-\mu)^2)$$

$= E[(X - E(X))^2]$

$= E[X^2 + (E(X))^2 - 2X E(X)]$

$= E[X^2] + E[(E(X))^2] - 2 E(X) E(X)$

$= E[X^2] + (E(X))^2 - 2(E(X))^2$

$= E[X^2] - [E(X)]^2$

Covariance: The covariance of 2 random variables is defined by

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - (E(Y))]$$

$$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$= E(XY) - E(Y).E(X) - E(Y)\cdot E(X) + E(E(X)\cdot E(Y))$$

$$= E(XY) - \cancel{E(Y).E(X)} - E(Y)\cdot E(X) + \cancel{E(X).E(Y)}$$

$$= E(XY) - E(Y)\cdot E(X)$$

So, from this,

* $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

* $\text{Cov}(X, X) = \text{Var}(X)$

* $\text{Cov}(X_1 + X_2, Y) = E[(X_1 + X_2)\cdot Y] - E(X_1 + X_2)\cdot E(Y)$

$$= E[X_1 \cdot Y] + E[X_2 \cdot Y] - E(Y)[E(X_1) + E(X_2)]$$

$$= \cancel{E(X_1 Y}$$

$$= \underline{E[X_1 \cdot Y]} + \underline{E[X_2 \cdot Y]} - \underline{E(Y)\cdot E(X_1)} - \underline{E(X_2)\cdot E(Y)}$$

$$= \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$$

In general, $\text{Cov}\left(\sum_{i=1}^{n} X_i, Y\right) = \sum_{i=1}^{n} \text{Cov}(X_i, Y)$

* $\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\left(\text{Cov}\left(X_i, \sum_{j=1}^{m} Y_j\right)\right)$

$$= \sum_{i=1}^{n} \text{Cov}\left(\sum_{j=1}^{m} Y_j, X_i\right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \text{Cov}(Y_j, X_i)$$

TYPES OF RANDOM VARIABLES (DISCRETE)

1. Bernoulli random variable:

Considering toss of a coin where

$p(\text{Head}) = p$

$p(\text{Tail}) = 1 - p.$

We define a Bernoulli random variable taking $X = 0$ and 1.

$$X = \begin{cases} 1 & \text{if } H \\ 0 & \text{if } T \end{cases}$$

PMF: $p_X(1) = P(X = 1) = p$ and $p_X(0) = P(X = 0) = 1 - p_X(1)$
$= 1 - p.$

So, $$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

$$= p_X(x) = p^x (1-p)^x \quad \text{for } x = 0, 1.$$

2. Binomial random variable:

Suppose there are n independent trials, each of which results in a success with probability p and failure with $1 - p$.

If X is a random variable with number of success in n trials,

$X \sim \text{Bin}(n, p)$.

PMF: $$p_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}. \quad x = 0, 1, 2, \ldots, n.$$

when $$\binom{n}{x} = {}^nC_x = \frac{\lfloor n}{\lfloor x \lfloor n-x}$$

3. Geometric random variable:

Let X be a random variable $=$ x trials till first success.

$$p_X(x) = p(X = x) = p(1-p)^{x-1}$$

4. Poisson random variable:

A better approximation of binomial variable (Limiting case of $X \sim \text{Bin}(n, p)$)

When n = very high, p = very small.

Let $X \sim \text{Bin}(n, p)$

Let $\lambda = np$.

$X \sim Bin(n, p)$

$\Rightarrow p_X(x) = P(X = x) = {}^nC_x\, p^x (1-p)^{n-x}$

$$\lim_{n\to\infty} {}^nC_x\, p^x (1-p)^{n-x}$$

$$\Rightarrow \lim_{n\to\infty} \frac{\lfloor n}{\lfloor x \lfloor n-x} p^x (1-p)^{n-x}$$

$$\Rightarrow \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-(x-1))\cancel{\lfloor n-x}}{\lfloor x \cancel{\lfloor n-x}} p^x (1-p)^{n-x}$$

$$\Rightarrow \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots\{n-(x-1)\}}{\lfloor x} \frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$\Rightarrow \lim_{n\to\infty} \frac{\cancel{n^x}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{\lfloor x} \frac{\lambda^x}{\cancel{n^x}}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \lambda^x \lim_{n\to\infty} \frac{(1-1/n)(1-2/n)\cdots\left(1-\frac{x-1}{n}\right)}{\lfloor x} \cdot \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{\lfloor x} \cdot \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} \cdot \left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{\lfloor x} e^{-\lambda}.$$

$$\therefore p_X(x) = \frac{\lambda^x}{\lfloor x} e^{-\lambda}.$$

5. <u>Pascal Random Variable</u>:

An extension of the geometric random variable.

X = Number of trials until Kth success.

P.M.F: $X \sim Pascal(n, p)$

~~$p_X(x)$~~

$$p_X(k) = P(X = k) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

ii) Uniform random variable:

$X \sim Unif(\{x_1, x_2, x_3, \ldots x_n\})$

$p_X(x_k) = P(X = x_k) = 1/n$.

Similarly for any definite interval $[a, b]$

$p_X(x_k) = P(X = x_k) = \frac{1}{b-a}$.

Expected value of all discrete random variables:

1. Bernoulli:

$$p_X(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \end{cases}$$

$$E(X) = \sum x\, p_X(x)$$
$$= 0(1-p) + 1 \cdot p = p.$$

2. Binomial:

$$p_X(x) = {}^nC_x\, p^x (1-p)^{n-x}$$

$$E(X) = \sum x\, p_X(x)$$
$$= \sum_{i=0}^{n} x \cdot {}^nC_x\, p^x (1-p)^{n-x}$$
$$= \sum_{i=0}^{n} \frac{\lfloor n \cdot x}{\lfloor x \lfloor n-x} p^x (1-p)^{n-x}$$
$$= np \sum_{i=1}^{n} \frac{\lfloor n-1}{\lfloor x-1 \lfloor (n-1)-(x-1)} p^{(x-1)} (1-p)^{(n-1)-(x-1)}$$
$$= np\,[p + (1-p)]^{n-1}$$
$$= np.$$

3. Poisson:

$$p_X(x) = \frac{\lambda^x e^{-\lambda}}{\lfloor x} \qquad \text{given } \lambda = np.$$

$$E(X) = \sum_{i=0}^{n} x\, p_X(x)$$
$$= \sum_{i=1}^{n} x \frac{\lambda^x e^{-\lambda}}{\lfloor x} = \sum_{i=1}^{n} \frac{\lambda \cdot \lambda^{x-1}}{\lfloor x-1} e^{-\lambda} = \lambda e^{-\lambda} \sum_{i=1}^{n} \frac{\lambda^{x-1}}{\lfloor x-1}$$
$$= \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda.$$

4. <u>Geometric Random variable</u>

$$p_x(x) = p(1-p)^{x-1}$$

$$E(X) = \sum_{i=1} xp(1-p)^{x-1} \quad — ①$$

Multiply $(1-p)$ in ①

$$(1-p)E(X) = \sum_{i=1} xp(1-p)^{x}$$

Subtract ① by ②,

$$E(X) - (1-p)E(X) = \sum_{i=1} xp(1-p)^{x-1} - \sum_{i=1} xp(1-p)^{x}$$

$$= xp \sum_{i=1} \left[(1-p)^{x-1} - (1-p)^{x}\right]$$

$$= xp \sum_{i=1} (1-p)^{x} \left[\frac{1}{1-p} - 1\right]$$

$$= x$$

$$E(X) = 1/p.$$

5. <u>Pascal</u>:

$$p_x(x) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

Expected value $E(X) = \frac{K}{P}$

6. <u>Uniform</u>:

$$X \sim Unif(\{\overset{a}{a}, a+1, a+2, \dots, \overset{b}{a+n}\})$$

$$E(X) = \frac{a + a+1 + a+2 + \dots + a+n}{n+1}$$

$$= \frac{(n+1)a + n(n+1)/2}{n+1}$$

$$= a + \frac{n}{2} = \frac{2a+n}{2} = \frac{a+a+n}{2} = \frac{a+b}{2}$$

$$E(X) = \frac{a+b}{2}.$$

Variances:

1. Bernoulli: ~~Rat~~

$E(X) = p.$

$$\begin{aligned}Var(X) &= E(X^2) - (E(X))^2 \\ &= \sum_{x=0,1} x^2 p_X(x) - (p)^2 \\ &= 0^2(1-p) + 1^2 \cdot p - p^2 \\ &= p - p^2 = p(1-p).\end{aligned}$$

2. Binomial:

$E(X) = np$

$$\begin{aligned}E(X^2) &= \sum x^2 p_X(x) \\ &= \sum x^2 \cdot {}^nC_x \, p^x (1-p)^{n-x} \\ &= \sum x^2 \frac{\underline{\lfloor n}}{\underline{\lfloor x}\,\underline{\lfloor n-x}} p^x (1-p)^{n-x} \\ &= \sum [x + x(x-1)] \, p_X(x) \\ &= \sum x \, p_X(x) + \sum x(x-1) \, p_X(x) \\ &= np + \sum x(x-1) \frac{\underline{\lfloor n}}{\underline{\lfloor x}\,\underline{\lfloor n-x}} p^x (1-p)^{n-x} \\ &= np + \sum \frac{\underline{\lfloor n}}{\underline{\lfloor x-2}\,\underline{\lfloor n-x}} p^x (1-p)^{n-x} \\ &= np + n(n-1) \sum \frac{\underline{\lfloor n-2}}{\underline{\lfloor x-2}\,\underline{\lfloor (n-\cancel{2})-(x-2)}} p^2 \cdot p^{x-2} (1-p)^{(n-2)-(x-2)} \\ &= np + n(n-1) p^2 \sum \frac{\underline{\lfloor n-2}}{\underline{\lfloor x-2}\,\underline{\lfloor n-x}} p^{x-2} (1-p)^{(n-2)-(x-2)} \\ &= np + p^2 n(n-1) + (p + (1-p))^{n-2} \\ &= np + np^2(n-1)\end{aligned}$$

$$\begin{aligned}Var(X) &= E(X^2) - (E(X))^2 \\ &= np + np^2(n-1) - n^2p^2 \\ &= np + \cancel{n^2p^2} - np^2 \cancel{= n^2p^2} \\ &= np(1-p).\end{aligned}$$

Poisson:

$$E(X) = \lambda$$

$$E(X^2) = \sum x^2 p_X(x)$$

$$= \sum_{i=1} x^2 \frac{\lambda^x e^{-\lambda}}{\lfloor x}$$

$$= \sum_{i=1} x \frac{\lambda^x e^{-\lambda}}{\lfloor x-1} = e^{-\lambda} \sum_{i=1} \frac{\{(x-1)+1\} \lambda^x}{\lfloor x-1}$$

$$= e^{-\lambda} \sum_{i=1} \left\{ \frac{\lambda^x}{\lfloor x-2} + \frac{\lambda^x}{\lfloor x-1} \right\}$$

$$= e^{-\lambda} \sum_{i=2} \frac{\lambda^x}{\lfloor x-2} + \sum_{i=1} \frac{\lambda^x}{\lfloor x-1}$$

$$= e^{-\lambda} \left[\lambda^2 \sum_{i=2} \frac{\lambda^{x-2}}{\lfloor x-2} + \lambda \sum_{i=1} \frac{\lambda^{x-1}}{\lfloor x-1} \right]$$

$$= e^{-\lambda} \left[\lambda^2 e^{\lambda} + \lambda e^{\lambda} \right]$$

$$= \lambda^2 + \lambda.$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2} = \lambda.$$

4. Geometric:

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

5. Pascal:

$$E(X) = \frac{k}{p}.$$

$$Var(X) = \frac{k(1-p)}{p^2}$$

6. Uniform:

$$E(X) = \frac{a+b}{2}.$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{(b-a)^2}{12}$$