Theory of Probability:

Random experiment: An experiment is said to be random if its outcome cannot be predicted precisely because the condition under which it is performed can not be predetermined with sufficient accuracy and completeness.

e.g. Tossing a coin, rolling a die etc.

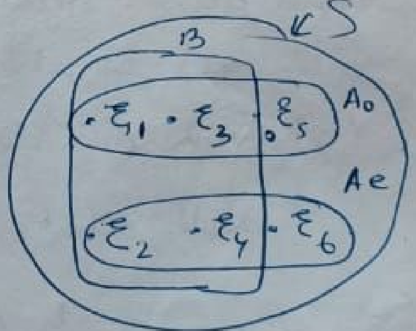
A random experiment may have several identifiable outcomes. e.g. rolling a die has six possible outcomes (1, 2, 3, 4, 5 and 6).

→ The sets of outcome meeting some specifications are known as events.

e.g. in rolling a die, the event "odd number on a throw" can result from any one of the three outcomes viz 1, 3 or 5. Hence, this event is a set of three outcomes. Thus, events are groupings of outcomes into classes among which we choose to distinguish.

We can have better idea understanding by using the concepts of set theory.

Sample space (S) is the collection of all possible separately identifiable outcomes of a random experiment. Each outcome is an element, or sample point, of this space and can be conveniently represented by a point in the sample space.



In random experiment of rolling a die, the sample space consists of six elements represented by six points $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ and ξ_6. The event on the other hand is a subset of S. The event an odd number is thrown, denoted by A_o is a subset of S. Similarly the event A_e an even number is thrown is another subset of S.

$A_o = \{\xi_1, \xi_3, \xi_5\}$. $A_e = (\xi_2, \xi_4, \xi_6)$

Let us denote the event "a number equal to or less than 4 is thrown" as B. Thus, $B = (\xi_1, \xi_2, \xi_3, \xi_4)$. Note that an outcome can also be an event, because an outcome is a subset of S with only one element.

The complement of any event A, denoted by A^c, is the event containing all points not in A. Thus for the event B, $B^c = (\xi_5, \xi_6)$, $A_o^c = A_e$ and $A_e^c = A_o$. An event that has no sample points is a null event, which is denoted by ϕ and equal to S^c.

The union of events A and B, denoted by $A \cup B$, is that event which contains all points in A and B. This is the event A or B.

$$A_o \cup B = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5)$$

$$A_e \cup B = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_6)$$

Note that $A \cup B = B \cup A$

The intersection of events A and B, denoted by $A \cap B$ or simply AB, is the event containing points common to A and B. This ~~the~~ is the event "both A and B" also known as the joint event AB. Thus the event $A_e B$, or "a number that is even and equal to or less than 4 is thrown," is a set (ξ_2, ξ_4) and similarly $A_o B = (\xi_1, \xi_3)$

Note that $AB = BA$

If the events A and B are such that

$AB = \phi$, then A and B are said to be disjoint or mutually exclusive events. This means events A and B cannot occur simultaneously. For example, A_o and A_e are mutually exclusive, meaning that in any ~~event~~ trial of the experiment if A_e occurs, A_o cannot occur at the same time, and vice versa.

Relative frequency and Probability:

Although the outcome of random experiment is unpredictable, there is statistical regularity about the outcomes. For example, if a coin is tossed a large number of times, about the half times the outcomes will be heads and the remaining half of the times will be tails. We say that the relative ~~density~~ frequency of the two outcomes heads or tails is one-half.

If A be one of the events of a random experiment. If we conduct a sequence of N independent trials of this experiment, and if the event A occurs in N(A) out of these N trials, then the fraction

$$f(A) = \lim_{N \to \infty} \frac{N(A)}{N}$$

is called the relative frequency of the event A.

Note that for small value of N, the fraction $N(A)/N$ may vary widely with N. As N increases, the fraction will approach a limit because of statistical regularity.

∴ The probality that an event A occurs is given by

$$P(A) = \lim_{N \to \infty} \frac{N(A)}{N}$$

We come to know that $0 \le P(A) \le 1$

Let us consider two events A and B of a random experiment. Suppose we conduct N independent trials of this experiment, and event A and B occurs in N(A) and N(B) trials respectively. If A and B are mutually exclusive (or disjoint), then if A occurs, B cannot occur and vice versa. Hence, the event

$A \cup B$ occurs in $N(A) + N(B)$ trials and

$$P(A \cup B) = \lim_{N \to \infty} \frac{N(A) + N(B)}{N}$$

$$= P(A) + P(B) \text{ if } AB = \phi$$

This result can be extended to more than two mutually exclusive events.

<u>Conditional Probability and Indipendent events</u>

<u>Conditional Probability</u>: We come across a situation where the probability of one event is influenced by the outcome of another event. e.g. consider drawing of two cards in succession from a deck. Let A denote the event the first card drawn is and an ace. We do not replace the card drawn in the first trial. Let B denote the event that the second card drawn is an ace. It is evident that the ~~drawing of~~ probability of drawing an ace in the second trial will be influenced by the outcome of the first draw. If the first draw does not result in an ace, the the probability of obtaining an ace in the second trial is $4/51$. The probability of event B thus depends on whether or not event A occurs.

So, the conditional probabity $P(B/A)$ [read as the probability of B given A) denotes the probabity of ~~B~~ event B when it is known that event A has ~~occu~~ occured.

4

Let an experiment be performed N times, in which the event A occurs n_1 times. Of these n_1 trials, event B occurs n_2 times. It is clear that n_2 is the number of times the joint event AB occurs. Thus.

$$P(BA) = \lim_{N \to \alpha} \frac{n_2}{N} = \lim_{N \to \alpha} \frac{n_1}{N} \frac{n_2}{n_1}$$

Note that $\lim_{N \to \alpha} \frac{n_1}{N} = P(A)$ and $\lim_{N \to \alpha} \frac{n_2}{n_1} = P(B/A)$ because B occurs n_2 times of n_1 times that A has occured. This represents the conditional probability of B given A

Therefore $P(BA) = P(A)\,P(B/A)$

$$\Rightarrow P(B/A) = \frac{P(BA)}{P(A)} \quad \text{provided } P(A) \neq 0.$$

Similarly $P(A/B) = \frac{P(AB)}{P(B)}$ provided $P(B) \neq 0$

From the two equations,

$$P(A/B) = \frac{P(A)\,P(B/A)}{P(B)}$$

and $$P(B/A) = \frac{P(B)\,P(A/B)}{P(A)}$$

The above two equations are called Bayes' rule. In Bayes' rule, one conditional probability is expressed in terms of the reverse conditional probability.

Example: A random experiment consists of drawing two cards from a deck in succesion without replacement. Find the probality of obtaining two red aces in two draws.

Solⁿ Let A and B the events "red ace in the first draw" and "red ace in the second draw". We wish to determine

$P(AB) = P(A)\,P(B/A)$

The relative frequency of A is $\frac{2}{52} = \frac{1}{26}$, Hence
$$P(A) = 1/26.$$

Also ~~the~~, $P(B/A)$ is the probability of drawing a red ace in the second draw given that the first draw was a red ace. The relative frequency is $1/51$, so
$$P(B/A) = 1/51.$$

$$\therefore P(AB) = P(A)\,P(B/A)$$
$$= \frac{1}{26} \times \frac{1}{51} = \frac{1}{1326}$$

Independent events: In such events, the occurance of one event in no way influences the occurance of other event. As an example, let us consider the drawing of two cards in succession, but in this case we replace the card obtained in the first draw and shuffle the deck before the second draw. In this case the outcome of the ~~first~~ second draw is in no way influenced by the outcome of the first draw.

The event B is said to be independent of the event A if
$$P(B/A) = P(B)$$

If the event B is independent of event A, then event A is also independent of B, i.e.
$$P(A/B) = P(A).$$

So, the probability that both events A and B occur is $P(AB) = P(A)\,P(B)$, * $\underline{P(A),\, P(B/A)}$

6

Bernoulli's Trial.

In Bernoulli Trials, if a certain event A occurs, we call it "success". If $P(A)=p$, then the probability of success is p. If q is the probability of failure, then $q = 1-p$.

Let us find the ~~probability of~~ probability of k successes in n (Bernoulli) Trials. The outcome of each trial is independent of the outcome of the other trials. It is clear that, if success occurs in k trials, the failure occurs in $n-k$ trials. Since the outcomes of the trials are independent, the probability of this event is clearly $p^k(1-p)^{n-k}$ i.e.

$$P(k \text{ successes in a specific order in } n \text{ trials}) = p^k(1-p)^{n-k}$$

But the event "k successes in n trials" can occur in many different ways. It is well known from combinational analysis that k things can be taken from n things in ${}^nC_k$ ways. where,

$$ {}^nC_k = \frac{n!}{k!(n-k)!}$$

This means the probability of k successes in n trials is

$$P(k \text{ successes in } n \text{ trials}) = {}^nC_k\, p^k(1-p)^{n-k}$$

$$= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

If we toss a coin n times, and the probability of observing head k times is given by

$$P(k \text{ heads in } n \text{ tosses}) = \frac{n!}{k!(n-k)!}(0{\cdot}5)^k(1-0{\cdot}5)^{n-k}$$

$$= \frac{n!}{k!(n-k)!}\, 0{\cdot}5^k \,.\, 0{\cdot}5^{n-k}$$

$$= \frac{n!}{k!(n-k)!}\, 0{\cdot}5^n$$

Examples.

A binary symmetric channel (BSC) has an error probability P_e (i.e. the probability of recieving 0 when 1 is transmitted, or vice versa is P_e). Note that the channel behaviour is symmetrical with respect to 0 and 1. Thus

$$P(0/1) = P(1/0) = P_e$$

$$P(0/0) = P(1/1) = 1 - P_e$$

where $P(y/x)$ denotes the probability of recieving y when x is transmitted. A sequence of n binary digits is transmitted over this channel. Determine the probability of recieving exactly k digits in error.

Sol^n The reception of each digit is independent of the other digits. This is an example of a Bernoulli trial with the probability of success $p = P_e$. ("success" here is recieving a digit in error). Clearly, the probability of k successes in n trials (k errors in n digits) is

$$P(\text{recieving } k \text{ errors out of } n \text{ digits}) = {}^nC_k \, P_e^k (1-P_e)^{n-k}$$

For example, if $P_e = 10^{-5}$, the probability of recieving two digits wrong in a sequence of eight digits is

$$^8C_2 (10^{-5})^2 (1-10^{-5})^6 \simeq \frac{8!}{2!\,6!} 10^{-10} = 2.8 \times 10^{-9}$$

Example- PCM Repeater Error Probability:

In PCM, regenerative repeaters are used to detect pulses, (before they are lost in noise) and retransmit new, clean pulses. This combats the accumulation of noise and pulse distortion.

A certain PCM channel consists of n identical links in tandem. The pulses are detected at the end of each link and clean new pulses are transmitted over the next link. If P_e is the probability of error in detecting a pulse over any one link, show that P_E, the probability of error in detecting a pulse over the entire channel is

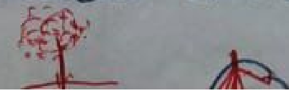
$$P_E \simeq nP_e \quad , \quad nP_e \ll 1 \checkmark$$

The probabilities of detecting a pulse correctly over one link and over the entire channel (n links in tadem) are $1-P_e$ and $1-P_E$, respectively. A pulse can be ~~to~~ detected correctly over the entire channel if either the pulse is detected correctly over every link or errors are made over an even number of links only

$$1-P_E = P(\text{correct detection over all links}) + P(\text{error over two links only}) + P(\text{error over four links only}) + \cdots + P(\text{error over } \alpha \text{ links only})$$

where α is n or n-1, depending ~~over~~ on ~~whether~~ whether n is even or odd. Because the pulse detection over each link is independent of the other links

$$P(\text{correct detection over all } n \text{ links}) = (1-P_e)^n$$



$$P(\text{error over } k \text{ links only}) = \frac{n!}{k!(n-k)!} P_e^k (1-P_e)^{n-k}$$

Hence, $$1 - P_E = (1-P_e)^n + \sum_{k=2,4,6..}^{\alpha} \frac{n!}{k!(n-k)!} P_e^k (1-P_e)^{n-k}$$

In practice, $P_e \ll 1$, so only the first two terms on the righthand side of this equation are of significance,

Also, $(1-P_e)^{n-k} \simeq 1$ and

$1-P_e \simeq 1$ if $P_e \ll 1$

$$1 - P_E \simeq (1-P_e)^n + \frac{n!}{2!(n-2)!} P_e^2 (1-P_e)^{n-2}$$

$$\simeq (1-P_e)^n + \frac{n(n-1)}{2} P_e^2 (1-P_e)^{n-2}$$

If $nP_e \ll 1$, then the second term can also be neglected, and

$$1 - P_E \simeq (1-P_e)^n$$

$$\Rightarrow 1 - P_E \simeq (1 - nP_e), \; nP_e \ll 1$$

and $P_E \simeq nP_E$

Q: In binary communication, one of the ~~techini~~ technique used to increase the reliability of a channel is to repeat a message several times. For example, we can send each ~~esar~~ message 0 or 1 three times. Hence the transmitted digits are 000 for message 0 or 111 for message 1. Because of channel noise, we may recieve any one of the eight possible combinations of three binary digits. The decission as to which the message is transmitted is made by the majority rule; that is, if at least two of the three detected digits are 0, the decission is 0 and so on. This scheme permits correct reception of data even if one out of three digits is in error. Detection error occurs only when if atleast two out of three digits are recived in error. If P_e is the error probability of one digit, and $P(\epsilon)$ is the probability of making a wrong decision in this scheme then

$$P(\epsilon) = {}^3C_2\, P_e^2 (1-P_e)^{3-2} + {}^3C_3\, P_e^3 (1-P_e)^{3-3}$$

$$= 3\, P_e^2 (1-P_e) + P_e^3$$

In practice $P_e \ll 1$, and

$$P(\epsilon) \simeq 3 P_e^2$$

For instance, if $P_e = 10^{-4}$, $P(\epsilon) = 3 \times 10^{-8}$. The error probability is reduced from 10^{-4} to 3×10^{-8}. We can use any odd number of repetitions for this scheme to function.

In this example, higher reliability is achieved at the cost of a reduction in the rate of information transmission by a factor of 3.

Random Variables:

The outcome of a random experiment may be a real e.g. rolling a die or it may be nonnumerical and describable by a phase such as heads or tails in tossing a coin. From a mathematical point of view, it is desirable to have numerical values ~~of~~ for all outcomes. For this reason, we assign a real number to each sample point according to some rule. If there are m sample points $\xi_1, \xi_2 \dots \xi_m$, then using some convinient rule, we assign a real number $x(\xi_1)$ to sample points $\xi_i (i = 1, 2, \dots, m)$. In case of tossing a coin, for example, we may assign the number 1 for the outcome heads and the number -1 for the outcome tails.

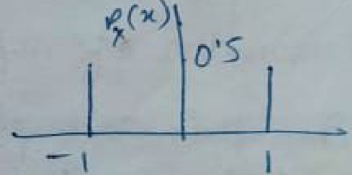
Thus, $x()$ is a function that maps sample points $\xi_1, \xi_2, \ldots \xi_m$ into real numbers $x_1, x_2, \ldots, x_m$. The variable X that takes on values $x_1, x_2, \ldots, x_n$ is called random variable.

We will use 'X' to denote random variable, and x_i to denote the value it takes. The probability of a random variable X taking a value x_i is $P_X(x_i)$

<u>Descrete Random variable:</u>

A random variable is descrete if there exists a denumerable sequence of distinct numbers x_i such that

$$\sum_i P_X(x_i) = 1$$



Thus a discrete RV can assume only a certain discrete values. An RV that can assume any value from a continuous interval is called <u>continuous random variable</u>.

We can ~~extend~~ extend the previous discussion to two RVs X and Y. <u>The joint probability</u> $P_{XY}(x_i, y_j)$ is the probability that $X = x_i$ and $Y = y_j$. Consider, the case of a coin two times in succession. If the outcomes of the first and second tosses are mapped into RV's X and Y, then X and Y each take values 1 and -1. Because the outcomes of the two tosses are independent, X and Y are independent and

$$P_{XY}(x_i, y_j) = P_X(x_i)\, P_Y(y_j)$$

and $P_{XY}(1,1) = P_{XY}(1,-1) = P_{XY}(-1,1) = P_{XY}(-1,-1) = 1/4$

6

These probabilities are plotted in the figure

For a general case where the variable x can take values $x_1, x_2, \ldots x_n$ and the variables ~~y~~ Y can take values $y_1, y_2, \ldots, y_m$ and we have

$$\sum_{i} \sum_{j} P_{XY}(x_i, y_j) = 1$$

Condi~~mati~~onal Probabilities.

If x and y are two Random Variables, the the conditional probability of $X = x_i$ and $Y = y_j$ is denoted by $P_{X/Y}(x_i/y_j)$.

Moreover $\sum_i P_{X/Y}(x_i/y_j) = \sum_j P_{Y/X}(y_j/x_i) = 1$.

Hence, $\sum_i P_{X/Y}(x_i/y_j)$ is the probability of the union of all possible ~~and~~ outcomes of x under given ~~dono~~ condition $Y = y_j$ and must be unity. A similar argument applies to $\sum_j P_{Y/X}(y_j/x_i)$

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

$$P_{XY}(x_i, y_j) = P_X(x_i) P_{Y/X}(y_j/x_i) = P_Y(y_j) P_{X/Y}(x_i/y_j)$$

Applying Bayes' rule $\left[P(A/B) = \frac{P(A) P(B/A)}{P(B)}\right]$

$$\sum_i P_{XY}(x_i, y_j) = \sum_i P_{X/Y}(x_i/y_j) P_Y(y_j)$$

$$= P_Y(y_j) \sum_i P_{X/Y}(x_i/y_j)$$

$$= P_Y(y_j)$$

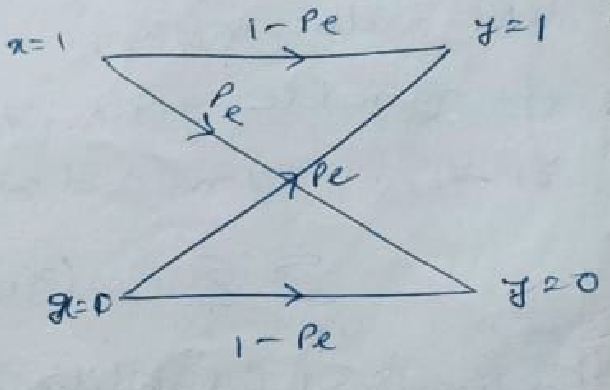
Similarly, $P_X(x_i) = \sum_j P_{XY}(x_i, y_j)$

The probabilities $P_X(x_i)$ and $P_Y(y_j)$ are called marginal probabilities. And the above equations shows how to find marginal probabilities from joint probabilities.

Ex

1. A binary symmetric channel (BSC) error probability is P_e. The probability of transmitting 1 is Q and that of transmitting 0 is 1 - Q. Determine the probabilities of receiving 1 and 0 at the receiver.

Sol^n If x and y are the transmitted and received digit, respectively, then for a BSC



$$P_{Y/X}(0/1) = P_{Y/X}(1/0) = P_e$$

$$P_{Y/X}(0/0) = P_{Y/X}(1/1) = 1 - P_e$$

Also, $P_X(1) = Q$ and $P_X(0) = 1 - Q$.

We need to find $P_Y(1)$ and $P_Y(0)$.

We have,

$$P_Y(y_j) = \sum_i P_{XY}(x_i, y_j)$$

$$= \sum_i P_X(x_i)\, P_{Y/X}(y_j/x_i)$$

$$\Rightarrow P_Y(1) = P_X(0)\, P_{Y/X}(1/0) + P_X(1)\, P_{Y/X}(1/1)$$

$$= (1-Q)P_e + Q(1-P_e)$$

Similarly,

$$P_Y(0) = P_X(0)\, P_{Y/X}(0/0) + P_X(1)\, P_{Y/X}(0/1)$$

$$= (1-Q)(1-P_e) + Q P_e$$

2

2 Over a certain binary communication channel, the symbol 0 is transmitted with probability 0.4 and 1 is transmitted with probability 0.6. It is given that $P(\epsilon/0) = 10^{-6}$ and $P(\epsilon/1) = 10^{-4}$, where $P(\epsilon/x_i)$ is the probability of detecting the error given that x_i is transmitted. Determine $P(\epsilon)$, the error probability of the channel.

Solⁿ If $P(\epsilon, x_i)$ is the joint probability that x_i is transmitted and it is detected wrongly, then the eqⁿ

$$P_x(x_i) = \sum_j P_{XY}(x_i, y_j) \text{ becomes,}$$

$$P(\epsilon) = \sum_i P(\epsilon, x_i) \cancel{\sum_i P(x_i, \epsilon)}.$$

$$= P(\epsilon, 0) + P(\epsilon, 1)$$

$$= P_x(0)\, P(\epsilon/0) + P_x(1)\, P_x(\epsilon/1)$$

$$= 0.4 \times 10^{-6} + 0.6 * 10^{-4}$$

$$= 0.604 \times 10^{-4}.$$

Note: $P(\epsilon/0) = 10^{-6}$ means that on the average, one out of 1 million received 0's will be detected erroneously. Similarly $P(\epsilon/1)$ means that on the average, one out of 10,000 received 1's will be in error.

But $P(\epsilon) = 0.604 \times 10^{-4}$ indicates that on the average, one out of $1/(0.604 \times 10^{-4}) \approx 16{,}566$ digits regardless of whether they are 1's or 0's will be received in error.

3

Probability models

Here we discuss two discrete functions, binomial and Poisson and two continuous functions ~~gau~~ gaussian and Rayleigh.

Biomomial Distribution: The binomial model describes an integer valued discrete RV associated with repeated trials.

A binomial random variable I corresponds to the number of times an event with probability α occurs in n ~~of~~ independent trials. ~~of an~~.

Thus this model applies to repeated coin tossing when I stands for the number of heads in n tosses and $P(H)=\alpha$. But more significantly, for us, it also applies to digital transmission when I stands for the number of errors in n-digit message with per digit error α probability α.

To formulate the binomial frequency function $P_I(i)=P(I=i)$, consider any sequence of n independent trials in which event A occurs ~~A~~ i times. If $P(A)=\alpha$ then $P(A^c)=1-\alpha$ and the sequence probability is $\alpha^i(1-\alpha)^{n-i}$. The member of different sequences with ~~I~~ i occurance is given by the binomial coefficient ${}^nC_i$, so we have

$$P_I(i) = {}^nC_i\, \alpha^i (1-\alpha)^{n-i}, \quad i=0,1,2,\dots n.$$

mean, $m = n\alpha$

variance $\sigma^2 = m(1-\alpha)$

4

Poisson Distribution.

The Poisson model also describes another integer-valued RV associated with repeated trials.

A ~~poi.~~ Poisson random variable I corresponds to the number of times an event occurs in an interval T when the probability of a single occurance in the small interval ΔT is $\mu \Delta T$.

The Poisson ~~distri~~ frequency function is

$$P_I(i) = e^{-\mu T} \frac{(\mu T)^i}{i!} \quad , \text{mean } m = \mu T$$
$$\text{variance } \sigma^2 = m$$

These expression describes random phenomenon such as radioactive decay and shot noise in electronic devices.

The Poission model is the approximation of binomial model when n is very large and α is very ~~no~~ small, and the product $n\alpha$ is finite. Under this condition the binomial model becomes awkward to compute.

Gaussian Probability density function:

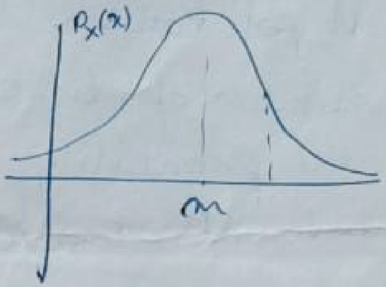
The Gaussian model describes a continuous RV having a normal distribution.

If x represents the sum of N independent random components, and if each component makes only a small contribution to the sum, then the cumulative distribution function of X approaches a gaussian CDF as N becomes large regardless of the distribution of the individual components.

5

A gaussian RV is a continuous random variable X with mean m, variance σ^2 and PDF

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$



This function is bell-shaped curve

The even symmetry about the peak at $x = m$ indicates that

$$P(X \le m) = P(X > m) = 1/2$$

Observe that X are just likely to fall above or below the mean.

If we know the mean m and variance σ^2 of the gaussian RV, then we can find the probability of the event $X > m + k\sigma$ by the following expression

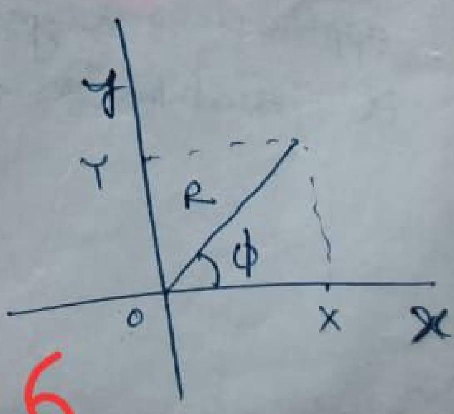
$$Q(k) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_k^{\infty} e^{-\lambda^2/2} d\lambda \quad \text{when } \lambda = \frac{x-m}{\sigma}$$

$$= P(X > m + k\sigma)$$

<u>Rayleigh PDF</u>

If X and Y are independent gaussian RVs with zero mean and same variance σ^2, the random variable be defined by $R = \sqrt{X^2 + Y^2}$ has a Rayleigh distribution.

To derive the corresponding Rayleigh PDF, we introduce the random angle ϕ and start with the joint PDF relationship



$$P_{R\phi}(r, \phi)\,|dr\,d\phi| = P_{XY}(x, y)\,|dx\,dy|$$

where $r^2 = x^2 + y^2$, $\phi = \tan^{-1}(y/x)$ $dx\,dy = r\,dr\,d\phi$

Since X and Y are independent gaussian with $m = 0$ and variance α^2

$$P_{XY}(x,y) = P_X(x) P_Y(y) = \frac{1}{2\pi\alpha^2} e^{-\frac{x^2+y^2}{2\alpha^2}}$$ (gaussian PDF)

$$P_X(x) = \frac{1}{\sqrt{2\pi\alpha^2}} e^{-\frac{(x-m)^2}{2\alpha^2}}$$

Since $r \geq 0$, we include $u(r)$

$$P_{R\phi}(r,\phi) = \frac{\cancel{r}\,1}{2\pi\alpha^2} e^{-r^2/2\alpha^2} u(r) \quad -① $$

The angle ϕ does not appear explicitly here, but its range is clearly limitted to 2π radians.

Integrating ① w.r.t. ϕ over the interval 0 to 2π we have

$$P_R(r) = \frac{r}{\alpha^2} e^{-r^2/2\alpha^2} u(r)$$

The mean is $\bar{R} = \sqrt{\frac{\pi}{2}}\,\alpha$ and second moment of R is $\bar{R^2} = 2\alpha^2$

Similarly we get the marginal probability for random angle ϕ as

$$P_\phi(\phi) = \int_0^\infty P_{R\phi}(r,\phi)\,dr = \frac{1}{2\pi}$$

Since R and ϕ are independent

$$P_{R\phi}(r,\phi) = P_R(r) P_\phi(\phi).$$

This result is used in the representation of bandpass noise.

Statistical Averages

The expectation (or mean) of a r.v. X, denoted by $E(x)$ is defined by

$$E(x) = \begin{cases} \sum_i x_i P_x(x_i) & x: \text{discrete} \\ \int_{-\infty}^{\infty} x P_x(x)\, dx, & x: \text{continuous}. \end{cases}$$

Variance.

The variance of a r.v. X, denoted by σ_x^2 or $Var(X)$ is defined as

$$Var(X) = \sigma_x^2 = E(X-m)^2$$

$$\therefore \quad \sigma_x^2 = \begin{cases} \sum_i (x_i - m)^2 P_x(x_i) & x: \text{discrete} \\ \int_{-\infty}^{\infty} (x-m)^2 P_x(x)\, dx, & x: \text{continuous}. \end{cases}$$

The positive square root of the variance, or σ_x, is called 'standard deviation' of X.

The variance or standard variation is a measure of the 'spread' of the values of X from its mean m.

$$\sigma_x^2 = E[X^2] - m^2 = E[X^2] - [E[X]]^2$$