ECE594I Notes set 2: Math

Mark Rodwell

University of California, Santa Barbara

rodwell@ece.ucsb.edu   805-893-3244, 805-893-3262 fax

Topics

Sets, probability, conditional probability, Bayes theorem, independence.

Bernoulli trials, random variables, density and distribution functions

Binomial distributions. Gaussian and Poisson as limiting cases.

Transformations.

Set Definitions

$\text{Set} = \text{ a collection of objects...}$

$\text{e.g.} \quad R = \{\text{set of all real \#s}\}$

$C = \{\text{set of cards in a deck}\}$

$\text{notation : elements lowercase, sets uppercase}$

$A = \{a, b, c, d\}$

$\text{Finite set : } \quad A = \{1,2,3,4\}$

$\text{Countably infinite set : } \quad B = \{0, 2, 4, 6, 8, ...\}$

$\text{uncountably infinite set : } \quad C = \{\text{real \#s between 1 and 3}\}$

Set Definitions

Subsets : $A \subseteq B$

A is a subset of B if all elements of A are also elements of B.

Strict subsets : $A \subset B$

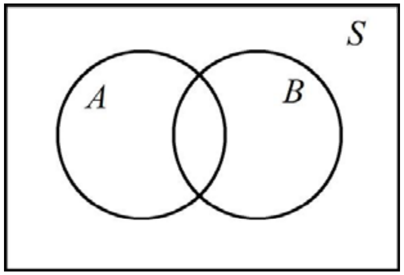
A is a strict subset of B

If all elements of A are also elements of B,

and if additionally B has elements which are not in A.

Set operations, logic operations

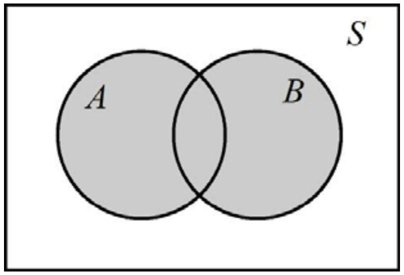
Sets A and B, both subsets of universal set S.



Union of Sets

$C = A \cup B$: set of all elements within either A or B.

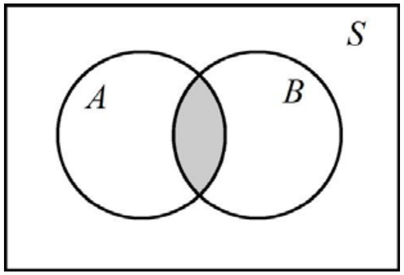
Boolean logic terminology: "A or B" $C = A + B$.



Intersection of Sets

$C = A \cap B$: set of all elements within both A or B.

Boolean logic terminology: "A and B" $C = AB$.



Other set operations

We can also define other set operations (and hence logic operations)

difference of sets

complements of sets

etc.,

...but we will asssume that the student knows these,

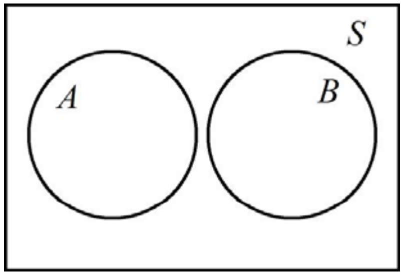
or can figure out as necessary.

Exclusive Sets

Key point for statistical independence.

Two sets are *exclusive* if $A \cap B = \{\ \}$ (the empty set),

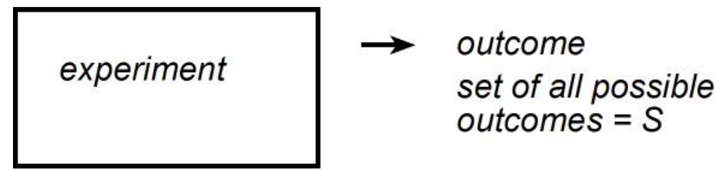
i.e. they have no common elements.



In probability, the events A and B are said to be mutually exclusive.

Probability

Key to the idea of probability is an *experiment*, for which there are a set of possible *outcomes*.



Each possible outcome has a numerical probability.

The sum of the probabilities of all *distinct* outcomes is one (something always happens).

Examples of probability

Toss a coin :

$$S=\{\text{heads, tails}\}=\{H,T\}$$

$$P(H)=P(T)=1/2$$

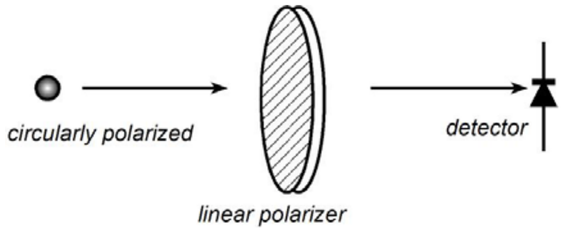
$$P(S)=P(H)+P(T)=1$$

Shoot a photon through a polarizer :

$$S=\{\text{photon received, no photon received}\}=\{R,N\}$$

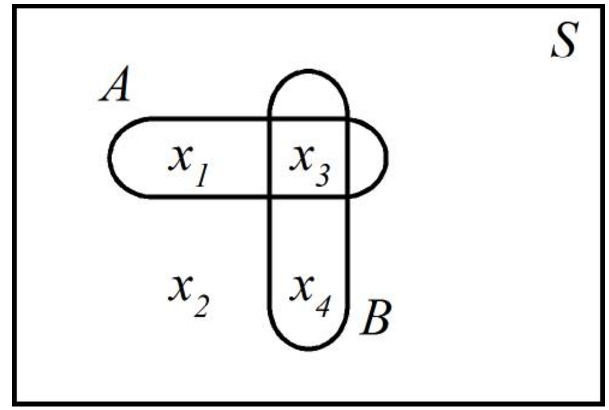
$$P(R)=P(N)=1/2$$

$$P(S)=P(R)+P(N)=1$$



Events

An event is a set of possible outcomes.



x_1, x_2, x_3, x_4 are possible outcomes.

$A = x_1 \cup x_2$ and $B = x_2 \cup x_4$ are events.

Rules (Axioms) of Probability

Each outcome has probability between 0 and 1.

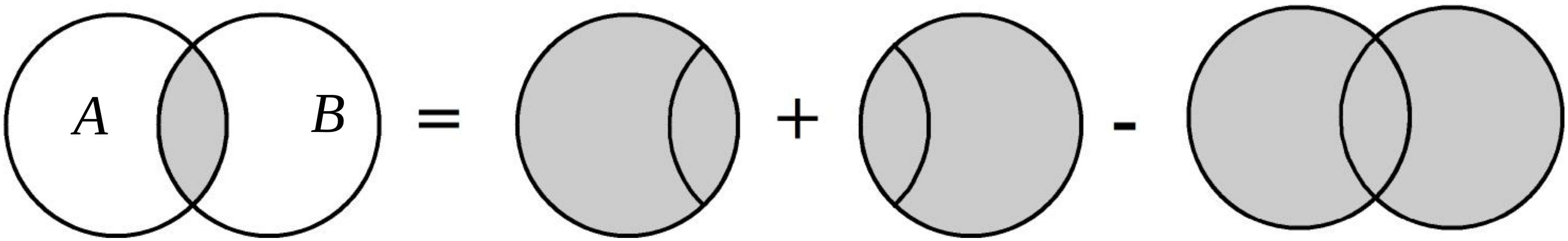
Given an event which is the union of muntually exclusive events, its probability is the sum of the individual probabilities.

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

if A, B, and C are distinct events.

Joint Probability

Probability that both A and B occur.



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

Conditional Probability

$P(A\,|\,B)$: probability of an event A, given the event B.

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

Total probability is obtained by adding up conditional probabilities:

$$P(A)=P(A\,|\,B_1)+P(A\,|\,B_2)+...+P(A\,|\,B_N)$$

But only if the events $B_1, B_2, ..., B_N$,

...are mutually exclusive, and together make up the sample space S.

i.e. $B_i \cap B_j = \{\ \}$ for $i \neq j$ and $B_1 \cup B_2 \cup ... \cup B_N = S$

Bayes' Theorem

Handy for receiver problems.

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

This follows directly from the definition of conditional probability:

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

Bayes' Theorem: Binary Communications Channel (1)

$$P(t=0)=T_0=1/2$$
$$P(t=1)=T_1=1/2$$

$$\text{transmitter} \xrightarrow{t} \text{channel} \xrightarrow{r} \text{receiver}$$

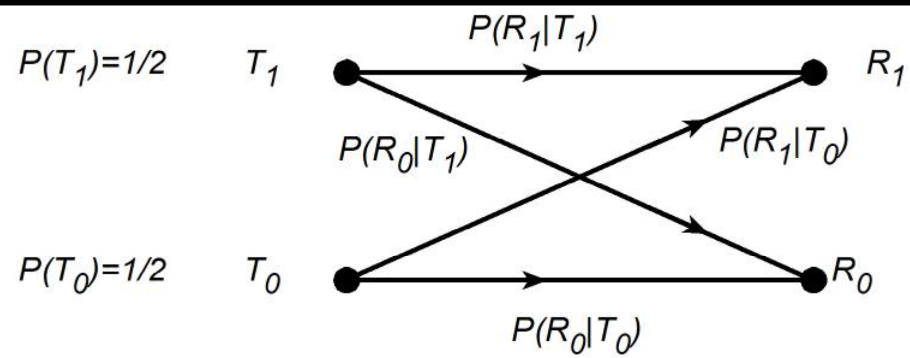
But the channel is noisy :

$$\left.\begin{aligned} P(r=0\,|\,t=0)&=1 \\ P(r=1\,|\,t=0)&=0 \end{aligned}\right\} \quad \text{The channel never make a mistake when a zero is sent.}$$

$$\left.\begin{aligned} P(r=0\,|\,t=1)&=1/2 \\ P(r=1\,|\,t=1)&=1/2 \end{aligned}\right\} \quad \text{..but makes mistakes when a one is sent.}$$

Bayes' Theorem: Binary Communications Channel (2)

Diagram



$$P(R_0 \mid T_0) = 1, \quad P(R_1 \mid T_0) = 0$$
$$P(R_0 \mid T_1) = 1/2, \quad P(R_1 \mid T_1) = 1/2$$

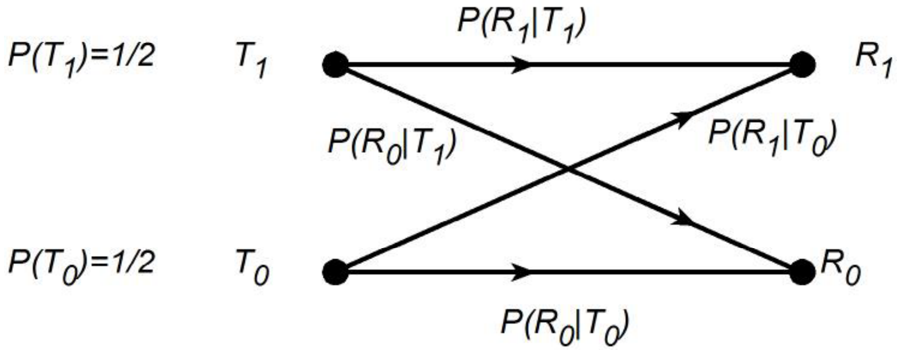
Probability that a zero is received :

$$P(R_0) = P(R_0 \mid T_0) \cdot P(T_0) + P(R_0 \mid T_1) \cdot P(T_1)$$
$$= (1) \cdot (1/2) + (1/2) \cdot (1/2) = 3/4$$

Bayes' Theorem: Binary Communications Channel (3)

$$P(R_0 \mid T_0) = 1, \quad P(R_1 \mid T_0) = 0$$

$$P(R_0 \mid T_1) = 1/2, \quad P(R_1 \mid T_1) = 1/2$$



If a zero is received, what is the probability that a one was sent ?

$$P(R_0 \mid T_1) = P(T_1 \mid R_0)\frac{P(T_1)}{P(R_0)} = \frac{1/2 * 1/2}{3/4} = 1/3$$

Before receiving the message, we know it has message has equal probabilities of being a 1 or a 0.

After receiving a zero, the odds are now 2/3 : 1/3 that a zero (vs. a 1) was sent.

Our knowledge has improved, but remains imperfect.

Independence

Two events, A and B, are statistically independent if their probabilities of occurence are unrelated.

i.e. $P(A \cap B) = P(A)P(B)$

recall : $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

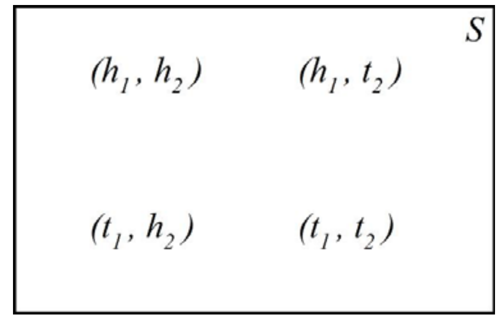
equivalently $P(A \mid B) = P(A)$

Independence of A and B is written as $A \perp\!\!\!\perp B$

Independence: Example

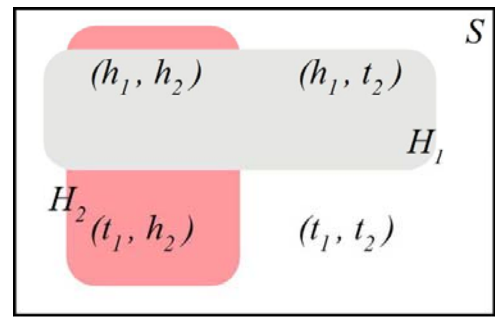
2 coin tosses : first toss h_1 or t_1, second toss h_2 or t_2

There are four outcomes



Define two events :

H_1 = first coin heads, $H_2 = 2^{nd}$ coin heads



Note that H_1 and H_2 are not mutually exclusive

If we assert from physical arguments that the 2 tosses are $\perp$,

then all 4 outcomes have 25% probability.

Indendence does not cascade...

If we have 3 events, A, B, and C,

then $A \perp\!\!\!\perp B$ and $B \perp\!\!\!\perp C$ does not imply $A \perp\!\!\!\perp C$

Bernoulli Trials→ Optical Shot Noise: Discrete-Time

Repeatedly peform a trial with binany outcomes.

Do the trial N times.

Each trial : probability of event A is P.

The number of sequences in which A occurs k times is

$$\# = \begin{pmatrix} N \\ k \end{pmatrix} = \text{"}N \text{ choose } k\text{"} = \frac{N!}{k!(N-k)!}$$

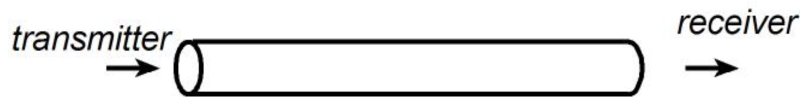
So the probability of k occurences in N trials is

$$P(k) = \begin{pmatrix} N \\ k \end{pmatrix} p^k (1-p)^{N-k} = \begin{pmatrix} N \\ k \end{pmatrix} p^k q^{N-k}$$

$q = (1-p)$ is a standard notational shortcut.

Bernoulli Trials→ Optical Shot Noise: Discrete-Time

Fiber : probability of photon transmission $= P$.



Event T_1 :

Transmitter sends N_1 photons for a message "1", with probability $= 1/2$.

Event T_0 :

Transmitter sends N_0 photons for a message "0", with probability $= 1/2$.

Passage of each photon is a Bernoulli trial.

$$P(k \text{ photons received} \mid N_1 \text{ sent}) = P(k \mid N_1) = \binom{N_1}{k} p^k q^{N_1-k}$$

$$P(k \text{ photons received} \mid N_0 \text{ sent}) = P(k \mid N_0) = \binom{N_0}{k} p^k q^{N_0-k}$$

Bernoulli Trials→ Optical Shot Noise: Discrete-Time

Problem is closely related to shot - noise - limited optical links.

We are at the receiver. What rule might we use to best decide what message was sent ?

If you receive k photons, is it more likely that a "1" or "0" was sent ?

Bernoulli Trials→ Optical Shot Noise: Discrete-Time

$$P(k) = P(k \mid T_o)P(T_o) + P(k \mid T_1)P(T_1)$$

$$= \frac{1}{2}\binom{N_1}{k} p^k q^{N_1-k} + \frac{1}{2}\binom{N_0}{k} p^k q^{N_0-k}$$

So the probability that a "1" was sent given k photons received is

$$P(T_1 \mid k) = P(k \mid T_1)P(T_1)/P(k)$$

$$= \frac{\frac{1}{2}\binom{N_1}{k} p^k q^{N_1-k}}{\frac{1}{2}\binom{N_1}{k} p^k q^{N_1-k} + \frac{1}{2}\binom{N_0}{k} p^k q^{N_0-k}}$$

$$= \frac{1}{1 + \dfrac{N_0!(N_1-k)!}{N_1!(N_0-k)!} q^{N_0-N_1}}$$

Bernoulli Trials→ Optical Shot Noise: Discrete-Time

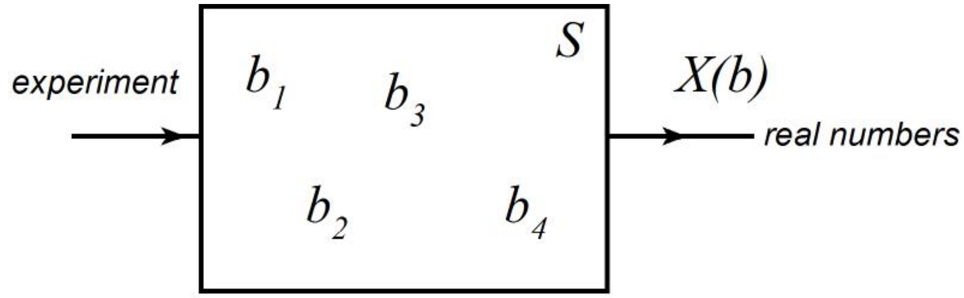
This at least tells us the relative odds of a "1" vs. a zero having been sent, given the message we have received.

Ultimately, we must still guess, with improved odds, as to the message sent.

One is tempted to choose the more probable outcome directly, but the best choice depends upon the *relative costs* of the types of possible errors in guessing.

Random Variables

Mathematician's picture :



outcomes b_i in sample space S

A random variable is a mapping of the sample space onto the set of real numbers.

Random variables must be bounded : $P\{X \to \pm\infty\} \to 0.$

We wish to define the probability that X falls within some range :
$P\{x_1 < X \le x_2\}.$

Discrete and Continuous Random Variables

Discrete R.V.'s

Only a discrete set of values: finite $\{1,2,3\}$, or infinite $\{1,2,3,4,\ldots\}$

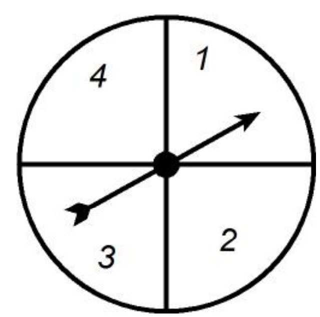
Continuous R.V.'s

Takes on a continuous range of values: example: $X \in \{\text{real \#s}\}$

We can have a continuous sample space and yet a discrete R.V.

Pointer: sample space $\Theta \in \{\text{real \#s between 0 and } 2\pi\}$

R.V. $\quad X \in \{1,2,3,4\}$



Cumulative Distribution Function

Random variable X.

Particular value it might take on : x.

Events : $\{X < x\}$ or $\{x_1 < X \le x_2\}$

The probability of the event $\{X < x\}$ is the cumulative distribution function

$$\underbrace{F_X(x)}_{} \equiv P(X \le x)$$

Distribution function of the random variable X.
at the particular outcome value x.

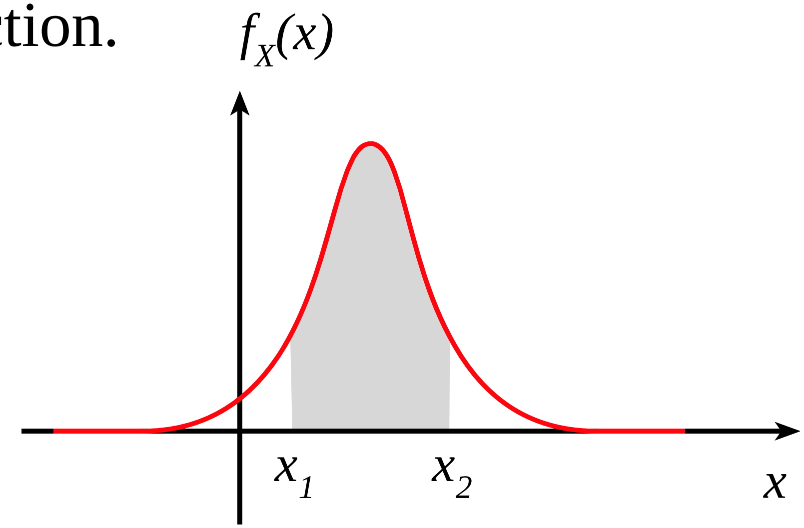
Probability Distribution Function

During an experiment, a random variable X takes on a particular value x.

The probability that x lies between x_1 and x_2 is

$$P\{x_1 < x < x_2\} = \int_{x_1}^{x_2} f_X(x)dx$$

$f_X(x)$ is the probability distribution function.

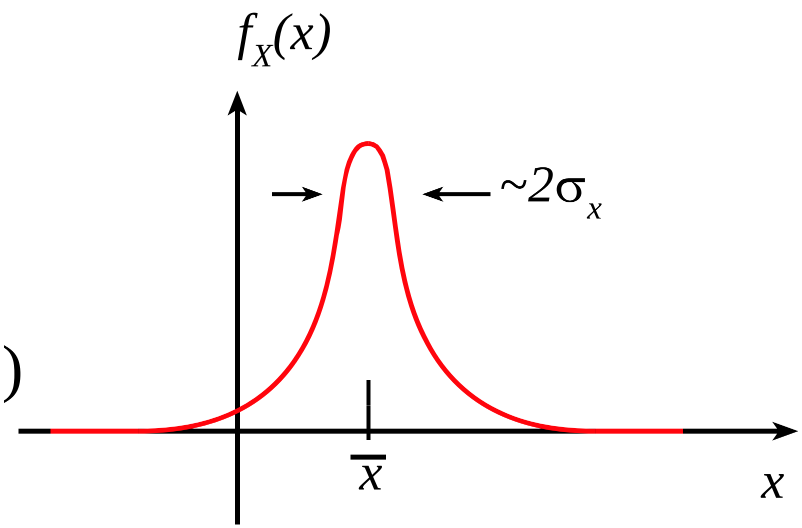


Example: The Gaussian Distribution

The Gaussian distribution:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(\frac{-(x-\bar{x})^2}{2\sigma_x^2}\right)$$

We will define shortly the mean ($\bar{x}$) and the standard deviation (σ_x^2).



Because of the *central limit theorem*, physical random processes arising from the sum of many small effects have probability distributions close to that of the Gaussian.

Gaussians also important because linear operations on Gaussians produce Gaussians $\rightarrow$ simplified math.

Gaussian Cumulative Distribution Function

$$F_x(x) = \int_{-\infty}^{x} f_X(x')dx' = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^{x} \exp\left(\frac{-(x'-\bar{x})^2}{2\sigma_x^2}\right)dx'$$

$$= \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^{\left(\frac{x-\bar{x}}{\sigma_x}\right)} \exp\left(\frac{-\beta^2}{2}\right)d\beta$$

There is no closed - form answer.

$$F_x(x) = 1 - Q\left(\frac{x'-\bar{x}}{\sigma_x}\right)$$

where $Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} \exp\left(\frac{-\beta^2}{2}\right)d\beta$

Error Functions

$Q(\alpha)$ can be related to the more well-known *error function*

$$Q(\alpha)=\frac{1}{2}\left[1-\operatorname{erf}\left(\frac{\alpha}{\sqrt{2}}\right)\right] \quad ,$$

but $Q(\alpha)$ is more directly useful in communcations problems.

I will provide a good tabulation of $Q(\alpha)$, but there is a very good bound for large α:

$$\left(\frac{\alpha^2}{1+\alpha^2}\right)\cdot\frac{1}{\alpha}\cdot\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\alpha^2}{2}\right)<Q(\alpha)<\frac{1}{\alpha}\cdot\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\alpha^2}{2}\right)$$

Tabulated values of the Q-function

ome values of the *Q*-functionare given below for reference.

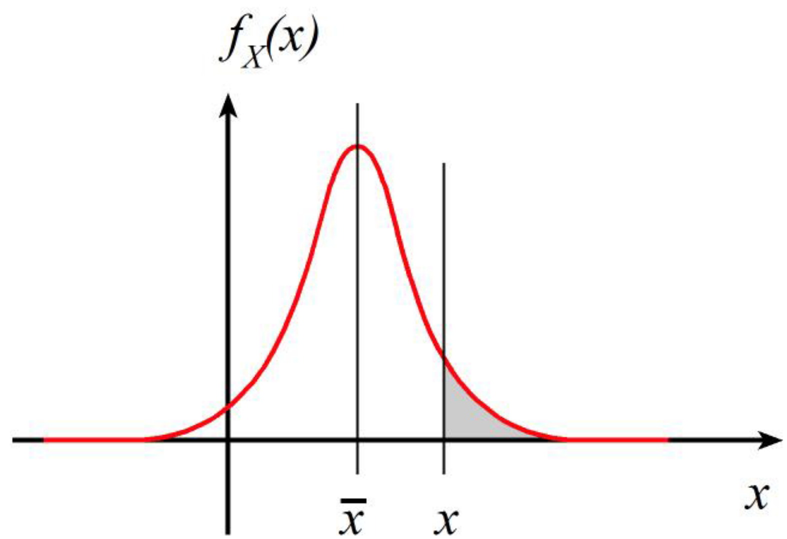
$Q(0.0)$ = 0.500000000	$Q(1.0)$ = 0.158655254	$Q(2.0)$ = 0.022750132	$Q(3.0)$ = 0.001349898
$Q(0.1)$ = 0.460172163	$Q(1.1)$ = 0.135666061	$Q(2.1)$ = 0.017864421	$Q(3.1)$ = 0.000967603
$Q(0.2)$ = 0.420740291	$Q(1.2)$ = 0.115069670	$Q(2.2)$ = 0.013903448	$Q(3.2)$ = 0.000687138
$Q(0.3)$ = 0.382088578	$Q(1.3)$ = 0.096800485	$Q(2.3)$ = 0.010724110	$Q(3.3)$ = 0.000483424
$Q(0.4)$ = 0.344578258	$Q(1.4)$ = 0.080756659	$Q(2.4)$ = 0.008197536	$Q(3.4)$ = 0.000336929
$Q(0.5)$ = 0.308537539	$Q(1.5)$ = 0.066807201	$Q(2.5)$ = 0.006209665	$Q(3.5)$ = 0.000232629
$Q(0.6)$ = 0.274253118	$Q(1.6)$ = 0.054799292	$Q(2.6)$ = 0.004661188	$Q(3.6)$ = 0.000159109
$Q(0.7)$ = 0.241963652	$Q(1.7)$ = 0.044565463	$Q(2.7)$ = 0.003466974	$Q(3.7)$ = 0.000107800
$Q(0.8)$ = 0.211855399	$Q(1.8)$ = 0.035930319	$Q(2.8)$ = 0.002555130	$Q(3.8)$ = 0.000072348
$Q(0.9)$ = 0.184060125	$Q(1.9)$ = 0.028716560	$Q(2.9)$ = 0.001865813	$Q(3.9)$ = 0.000048096
			$Q(4.0)$ = 0.000031671

http://en.wikipedia.org/wiki/Q-function

Error Functions

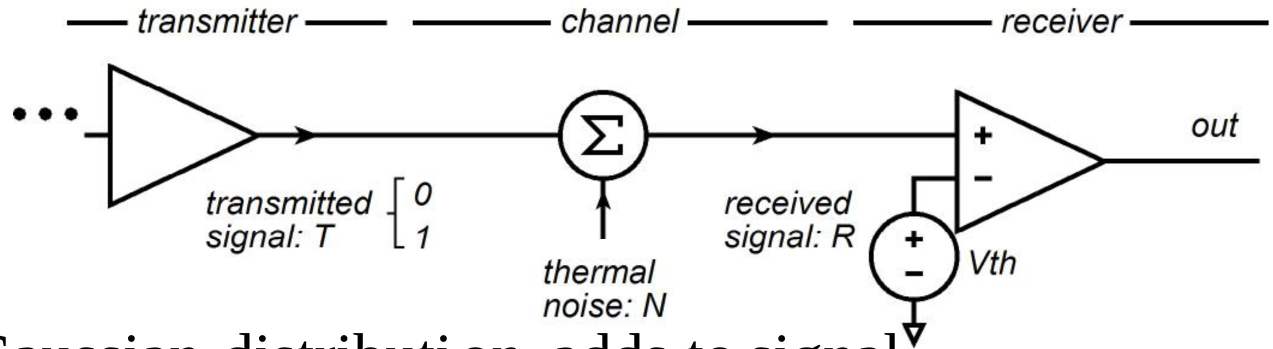
$$F_x(x)=\frac{1}{\sqrt{2\pi\sigma_x^2}}\int_{-\infty}^{x}\exp\left(\frac{-(x'-\bar{x})^2}{2\sigma_x^2}\right)dx'=1-Q\left(\frac{x'-\bar{x}}{\sigma_x}\right)$$

$$\text{where } Q(\alpha)=\frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\infty}\exp\left(\frac{-\beta^2}{2}\right)d\beta$$



$Q\left(\frac{x-\bar{x}}{\sigma_x}\right)$ gives the probability of the Gaussian exceeding its mean $(\bar{x})$ by $\left((x-\bar{x})/\sigma_x\right)$ standard deviations.

Example: Digital Transmission



N = thermal noise, Gaussian distribution, adds to signal.

Receiver decides between "0" & "1" depending on whether

$R = T + N$ is bigger than or smaller than 1/2.

$$f_T(t) = (1/2)\delta(t) + (1/2)\delta(t-1) \quad (t \text{ is transmitted signal, not time})$$

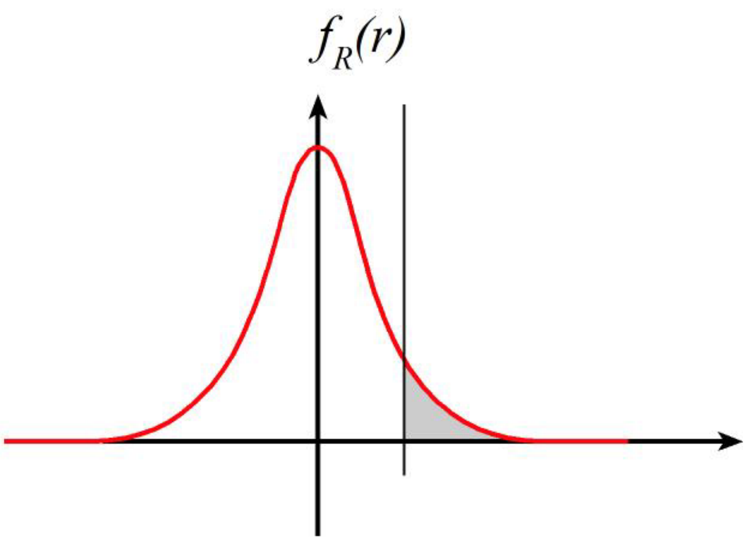
$$f_N(n) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(\frac{-\left(n-\overline{n}\right)^2}{2\sigma_n^2} \right) \text{ with } \overline{n} = 0 \text{ and } \sigma_n = 0.25 \text{ (say)}$$

Example: Digital Transmission

What is the probability of error, given that a zero was sent ?

$$P(\text{error} \mid T=0) = P(R > 1/2 \mid T=0) = P(N > 1/2\}$$
$$= 1 - P(N \le 1/2\} = 1 - F_N(n = 1/2)$$
$$= Q\left(\frac{1/2 - \overline{n}}{\sigma_N}\right) = Q(2)$$
$$= 0.023$$

In this context, the signal /noise ratio is 2 and $Q(2)$ is the probability of error.



Bernoulli Trials → Binomial Distribution

N Bernoulli trials :

$$P(k \text{ successes} \mid N \text{ trials}) = P(k \mid N) = \binom{N}{k} p^k q^{N-k}$$

note that k is a discrete random variable.

Call x the # of sucesses, a continuous random variable :

$$f_X(x) = \sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} \delta(x-k)$$

...which is just a change in notation.

Important because :

- arises in problems involving counting $\perp\!\!\!\perp$ events → shot noise
- leads to other distributions as limiting cases.

Binomial→ Gaussian

N Bernoulli trials :

$$P(k \text{ successes} \mid N \text{ trials}) = P(k \mid N) = \begin{pmatrix} N \\ k \end{pmatrix} p^k q^{N-k}$$

If (Papoulis 1965, p. 66) $npq >> 1$

and if $|k - np| \sim O\left(\sqrt{npq}\right)$ or less, then :

$$p_n(k) \cong \frac{1}{\sqrt{2\pi}\sqrt{npq}} \exp\left[-\frac{(k-np)^2}{2(npq)}\right]$$

This is a Gaussian of mean np and variance npq.

Binomial→ Poisson

$$N \text{ Bernoulli trials}: P(k \text{ successes} \mid N \text{ trials}) = P(k \mid N) = \binom{N}{k} p^k q^{N-k}$$

If (Papoulis 1965, p. 71) $n >> 1$, $p << 1$, but np is not $>> 1$, then :

$$\binom{N}{k} p^k q^{N-k} \cong e^{-Np} \cdot \frac{(Np)^k}{k!}$$

i.e. the distribution approaches the Possion distribution : $P(k) \cong e^{-a} \cdot \dfrac{a^k}{k!}$

Exponential Disribution

$$f_X(x) = \begin{cases} (1/a)\exp\left[-\dfrac{x-b}{a}\right] & \text{for } x > b \\ 0 & \text{otherwise} \end{cases}$$

This will show up in thermally - driven distributions (Boltzmann)

Transformations of a Random Variable

Suppose $Y = Y(X)$, where $Y(X)$ is a 1-1 function, and given some $f_X(x)$.

$P(y_1 < Y < y_2) = P(x_1 < X < x_2)$ where $y_1 = y(x_1), y_2 = y(x_2)$

If $x_2 = x_1 + \varepsilon$, then $y_2 = y_1 + \varepsilon \left.\frac{dy}{dx}\right|_{x_1}$

so,

$$\varepsilon \left.\frac{dy}{dx}\right|_{x_1} \cdot \left| f_Y(y_1) \right| = \varepsilon \cdot \left| f_X(x_1) \right|$$

or, more clearly

$$f_Y(y_1) = \frac{f_X(x_1)}{\left| \left.\frac{dy}{dx}\right|_{x_1} \right|}$$

Transformations of a Random Variable

This assumed a 1-1 function, i.e. each range in Y corresponds to a single range in X. This may not always be true...

Consider $y = cx^2$:

Then the regions $x \sim 1$ and x ~ -1 both map to y ~ c.

The density function for y is then

$$f_Y(y_1) = \left.\frac{f_X(x)}{2cx}\right|_{x=\sqrt{y/c}} + \left.\frac{f_X(x)}{2cx}\right|_{x=-\sqrt{y/c}} = \frac{f_X\left(\sqrt{y/c}\right) + f_X\left(-\sqrt{y/c}\right)}{2\sqrt{cy}}$$

References and Citations:

Sources / Citations :

Kittel and Kroemer : Thermal Physics

Van der Ziel : Noise in Solid - State Devices

Papoulis : Probability and Random Variables (hard, comprehensive)

Peyton Z. Peebles : Probability, Random Variables, Random Signal Principles (introductory)

Wozencraft & Jacobs : Principles of Communications Engineering.

Motchenbaker : Low Noise Electronic Design

Information theory lecture notes : Thomas Cover, Stanford, circa 1982

Probability lecture notes : Martin Hellman, Stanford, circa 1982

National Semiconductor Linear Applications Notes : Noise in circuits.

Suggested references for study.

Van der Ziel, Wozencraft & Jacobs, Peebles, Kittel and Kroemer

Papers by Fukui (device noise), Smith & Personik (optical receiver design)

National Semi. App. Notes (!)

Cover and Williams : Elements of Information Theory