Ministry of Higher Education and Scientific Research

Al-Furat Al- Awsat Technical University

Engineering Technical college / Najaf

Communication Engineering Department

Information Theory (CE231)

2nd Class 2018/2019

Lecturer Ali M. Alsahlany

Lecture Outlines :

- **Random Variables**
- **Discrete Probability Distribution**
- **Distribution Functions for Random Variables**
- **Expectation of a Discrete Random Variables**
- **Variance of a Discrete Random Variable**
- **Continuous Random Variables**
- **Distribution Functions for Continuous Random Variables**
- **Variance of a Discrete Random Variable**
- **The Binomial Distribution**
- **Normal distribution (Gaussian)**
- **Poisson Distributions**

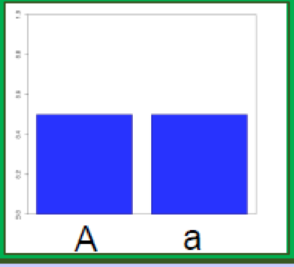
Random Variables: Suppose that to each point of a sample space we assign a number. We then have a function defined on the sample space. This function is called a random variable (or stochastic variable). It is usually denoted by a capital letter such as X or Y.

Example 1: X is the variable for the number of heads for a coin tossed three times

Solution: $X = 0, 1, 2, 3$

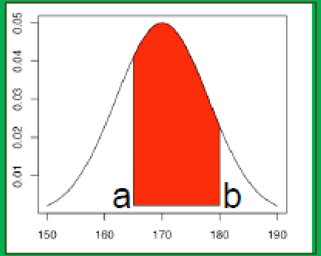
Discrete random variable has accountable number of possible values

- Number of sales, Number of calls, People in line, Mistakes per page, dice.



A continuous random variable takes all values in an interval of numbers

- electrical current, pressure, temperature, time, voltage, blood pressure, the speed of a car, the real numbers from 1 to 6.



The probability distribution of a discrete random variable is a graph, table or formula that specifies the probability associated with each possible outcome the random variable can assume.

- $f(x) \geq 0$ **for all values of x**
- $\sum_x f(x) = 1$ $f(x)$ **is the probability function**

Example 2: Suppose that a fair coin is tossed twice. Let *X* represent the number of heads that can come up. Find the probability function corresponding to the random variable X.

Solution: For each sample point we can associate a number for X as follows:

Sample Point	HH	HT	TH	TT
X	2	1	1	0

$P(HH) = 1/4,\ P(HT) = 1/4,\ P(TH) = 1/4,\ P(TT) = 1/4$

$P(X=0) = P(TT) = 1/4$

$P(X=1) = P(HT \cup TH) = P(HT)+P(TH)= 1/4+1/4 =1/2$

$P(X=2) = P(HH) = 1/4$

The probability function is given by

x	0	1	2
$f(x)$	1/4	1/2	1/4

Cumulative Distribution Function (CDF) F(x) **- is a function that returns the probability that a random variable X is less than or equal to a value. The CDF is also sometimes called the distribution function (DF).**

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

Requirements for CDFs

(1) $F(x) \geq 0$ **everywhere the distribution is defined**
(2) $F(x)$ **non-decreasing everywhere the distribution is defined.**
(3) $F(x) \to 1$ **as** $x \to \infty$

Example 3: Consider the probability distribution of the number of rewards you will get this semester

x	*f(x)*	*F(x)*
0	0.05	0.05
1	0.15	0.20
2	0.20	0.40
3	0.60	1.00

Expectation : **The expected value, or mean, of a random variable is a measure of central location.**

$$E(X) = \mu = \sum x.f(x)$$

Example 4: **Suppose that a game is to be played with a single die assumed fair. In this game a player wins 20\$ if a 2 turns up; 40\$ if a 4 turns up; loses 30\$ if a 6 turns up; while the player neither wins nor loses if any other face turns up. Find the expected sum of money to be won.**

Solution:

$f(x_1) = f(x_2) = f(x_3) = f(x_4) = f(x_5) = f(x_6) = 1/6$

x	0	+20	0	+40	0	-30
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

The expected value, or expectation, is

$$E(X) = (0)\left(\frac{1}{6}\right) + (20)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (40)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (-30)\left(\frac{1}{6}\right) = 5$$

The player should expect to pay 5 \$ in order to play the game.

Variance : ***A positive quantity that measures the spread of the distribution of the random variable about its mean value. Larger values of the variance indicate that the distribution is more spread out.***

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = \sum_{j=1}^{n}(x_j-\mu)^2 f(x_j) = \sum_x (x-\mu)^2 f(x)$$

The *standard deviation* is the positive square root of the variance

$$\sqrt{\sigma^2} = \sqrt{E[(X-\mu)^2]} = \sqrt{\sum (x-\mu)^2 f(x)}$$

Example 5: Find the variance and standard deviation for the game played in previous Example

Solution:

$E(X) = \square = 5$

x_j	0	+20	0	+40	0	-30
$f(x_j)$	1/6	1/6	1/6	1/6	1/6	1/6

$$\sigma_X^2 = (0-5)^2\left(\frac{1}{6}\right) + (20-5)^2\left(\frac{1}{6}\right) + (0-5)^2\left(\frac{1}{6}\right) + (40-5)^2\left(\frac{1}{6}\right) + (0-5)^2\left(\frac{1}{6}\right) + (-30-5)^2\left(\frac{1}{6}\right) = \frac{2750}{6}$$

$$\sigma^2 = \sqrt{\frac{2750}{6}}$$

Some properties of expected values :

- $E(a \cdot x) = a \cdot E(x) = a \cdot \mu$

where *a* is a constant

- $E(a \cdot x + b) = a \cdot E(x) + b = a \cdot \mu + b$ *where a and b are constants*
- $E(x + y) = E(x) + E(y)$

Some properties of variance :

- $\sigma^2 = E[(X-\mu)^2] = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2$ *where* $\mu = E(X)$
- $Var(cX) = c^2 Var(X)$ *where c is any constant*
- *The quantity* $E[(X-a)^2]$ *is a minimum when* $a = \mu = E(X)$
- *If X and Y are independent random variables,*

$Var(X+Y) = Var(X) + Var(Y)$ *or* $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$

$Var(X - Y) = Var(X) - Var(Y)$ *or* $\sigma^2_{X-Y} = \sigma^2_X - \sigma^2_Y$

Exercise 1: *compute the expected value and variance of the number of rewards in Example 3 in this Lecture.*

$E(x) = 0*.05+1*0.15+2*0.20+3*0.60 = 2.35$

$Var\ (x) = (0-2.35)^2 * 0.05+ (1-2.35)^2*0.15+(2-2.35)^2*0.20+(3-2.35)^2*0.60 = 0.8275$

Homework 1: *Let X and Y be the random independent events of rolling a fair die. Compute the expected value of X + Y, and the variance of X + Y.*

Homework 2: *The number of e-mail messages received per hour has the following distribution. compute the expected value and variance.*

x = number of message	10	11	12	13	14	15
f(x)	0.08	0.15	0.30	0.20	0.20	0.07

Continuous Random Variables :

A nondiscrete random variable *X* is said to be *absolutely continuous*, or simply *continuous*, if its distribution function may be represented as

$$F(x) = P\,(X \le x) = \int_{-\infty}^{x} f\,(u)du$$

where the function *f(x)* has the properties

1. $f(x) \ge 0$
2. $\int_{-\infty}^{\infty} f\,(x) = 1$

The function *f(x)* is called the probability density function (p.d.f.).

Example 6: **Find the constant c such that the function**

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & otherwise \end{cases}$$

is a density function, and then find $P(1 < X < 2)$.

Solution:

Notice that if $c \ge 0$, then Property 1 is satisfied. So $f\,(x)$ must satisfy Property 2 in order for it to be a density function.

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{3} cx^2\,dx = \left|\frac{cx^3}{3}\right|_0^3 = 9c$$

and since this must equal 1, c = $\frac{1}{9}$, and our density function is

$$f(x) = \begin{cases} \frac{1}{9}x^2 & 0 < x < 3 \\ 0 & otherwise \end{cases}$$

Next,

$$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

Cumulative Distribution Function (c.d.f.)

The c.d.f. of a continuous random variables is defined exactly the same as for discrete random variables

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

where x is any real number, i.e., $-\infty \leq x \leq \infty$. So,

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

Example 7: **Find the distribution function for example 6.**

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{9}x^2 dx = \frac{x^3}{27} \qquad \textit{where } x \leq 3.$$

Expectation :

If X is a continuous random variable having probability density function $f(x)$, then it can be shown that

$$\mu_x = E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Example 8: **The density function of a random X is given by**

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & otherwise \end{cases}$$

The expected value of X is then

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{0}^{2} x\left(\frac{1}{2}x\right) dx = \int_{0}^{2} \frac{x^2}{2}\, dx = \left.\frac{x^3}{6}\right|_0^2 = \frac{4}{3}$$

Variance :

If X is a continuous random variable having probability density function *f(x)*, then the variance is given by

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

Example 9: **Find the variance and standard deviation of the random variable from Example 8, using the fact that the mean was found**

To be $\mu_x = \text{E}(\text{X}) = \frac{4}{3}$

$$\sigma^2 = E\left[(X-\frac{4}{3})^2\right] = \int_{-\infty}^{\infty}\left(x-\frac{4}{3}\right)^2 f(x)dx = \int_{-\infty}^{\infty}\left(x-\frac{4}{3}\right)^2\left(\frac{1}{2}x\right)dx = \frac{2}{9}$$

And so the standard deviation is $\sigma = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}$.

Binomial: An experiment for which the following four conditions are satisfied is called a *binomial experiment*.

1. The experiment consists of a sequence of n trials, where n is fixed in advance of the experiment.
2. The trials are identical, and each trial can result in one of the same two possible outcomes, which are denoted by success (S) or failure (F).
3. The trials are independent.
4. The probability of success is constant from trial to trial: denoted by p.

The probability mass function (PMF) of X is

for x = 0; 1; 2; . . . ; n

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}$$

p = probability of success

q = *(1 – p)* probability of failure

n = number of trials

x = number of successes

Example 10: In a digital communication system, the number of bits in error in a packet depicts a Binomial discrete random variable

Example 11: The probability of getting exactly 2 heads in 6 tosses of a fair coin is

Solution:

$$P(X=2) = \binom{6}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{6-2} = \frac{6!}{2!\,4!}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4} = \frac{15}{64}$$

Exercise 2: If I toss a coin 20 times, what's the probability of getting 2 or fewer heads?

Solution:

$$\binom{20}{0}(.5)^{0}(.5)^{20} = \frac{20!}{20!0!}(.5)^{20} = 9.5x10^{-7} +$$

$$\binom{20}{1}(.5)^{1}(.5)^{19} = \frac{20!}{19!1!}(.5)^{20} = 20x9.5x10^{-7} = 1.9x10^{-5} +$$

$$\binom{20}{2}(.5)^{2}(.5)^{18} = \frac{20!}{18!2!}(.5)^{20} = 190x9.5x10^{-7} = 1.8x10^{-4}$$

$$= 1.8x10^{-4}$$

(Mean and Variance for Binomial Distribution) :

If X is a binomial random variable with parameters p and n, then

Mean $\mu = np$

Variance $\sigma^2 = npq$

Standard Deviation $\sigma = \sqrt{npq}$

Exercise 3: Digital Channel the chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X = the number of bits in error in the next four bits transmitted. Determine P(X = 2), the mean, variance, and standard deviation of this experiment.

Solution:

Outcome	x	Outcome	x
OOOO	0	*EOOO*	1
OOOE	1	*EOOE*	2
OOEO	1	*EOEO*	2
OOEE	2	*EOEE*	3
OEOO	1	*EEOO*	2
OEOE	2	*EEOE*	3
OEEO	2	*EEEO*	3
OOOO	3	*EEEE*	4

$$\mathbf{P(X = 2)} = \binom{4}{2}(0.1)^2(0.9)^{4-2}$$

$$\mathbf{E\,(X)} = \mu_x = 4(0.1) = 0.4 \qquad \sigma^2 = 4(0.1)(0.9) = 0.36$$

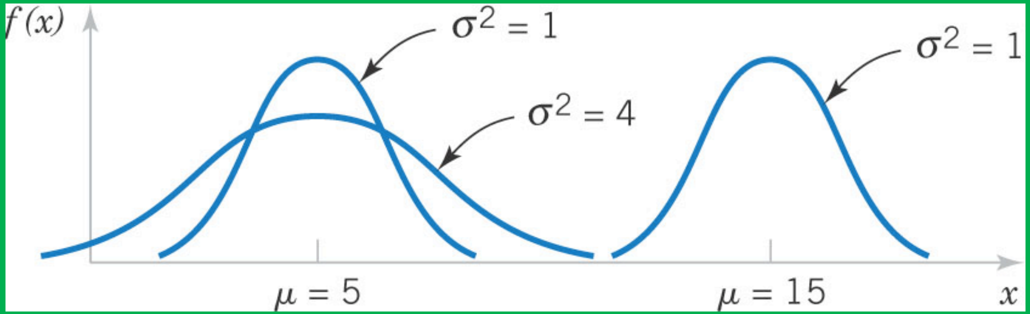
Homework 3: Suppose that three telephone users use the same number and that we are interested in estimating the probability that more than one will use it at the same time. If independence of telephone habit is assumed, the probability of exactly k persons requiring use of the telephone at the same time is given by the mass function $p_X(k)$ associated with the binomial distribution. Let it be given that, on average, a telephone user is on the phone 5 minutes per hour.

Homework 4: In a digital communication channel, assume that the number of bits received in error can be modeled by a binomial random variable. The probability that a bit is received in error is 10^{-5}. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur? Let X denote the number of errors.

Note : Clearly, this probability is difficult to compute. Fortunately, the normal distribution can be used to provide an excellent approximation in this example.

Homework 5: Consider the problem of missile firing. Enumerate the possible outcomes of trail firing of missiles. Lets S denote the success of each trail with probability *p* and *F* denote the failure with probability *q=1-p*, then there are, 2^4 possible outcomes s listed below as set U, the complete sample space.

- **One of the most important examples of a continuous probability distribution is the *normal distribution*, sometimes called the *Gaussian distribution*. most popular to communication engineers is … AWGN Channels.**
- **Random variation of many physical measurements are normally distributed.**
- **The location and spread of the normal are independently determined by mean (μ) and standard deviation (σ).**

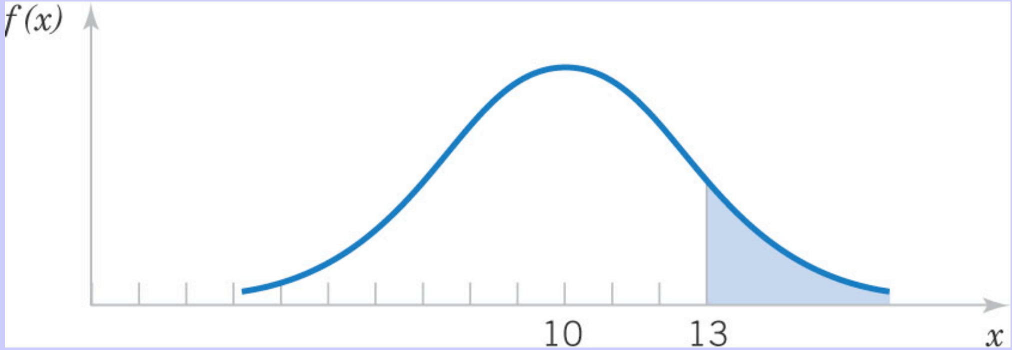


The density function for this distribution is given by

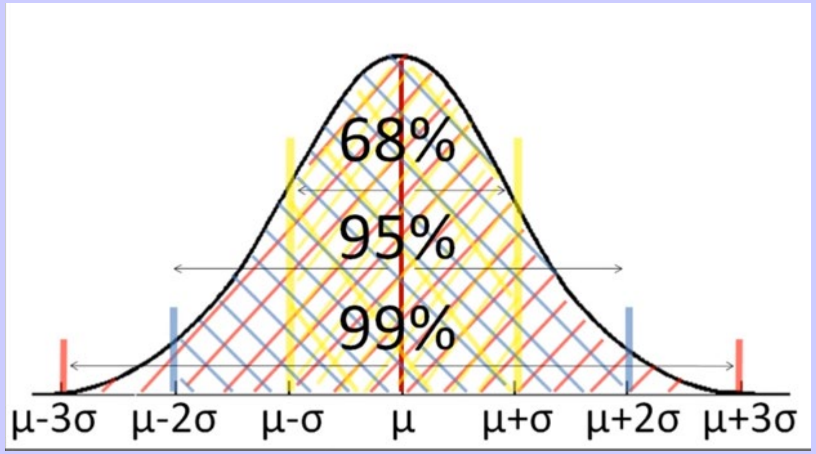
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

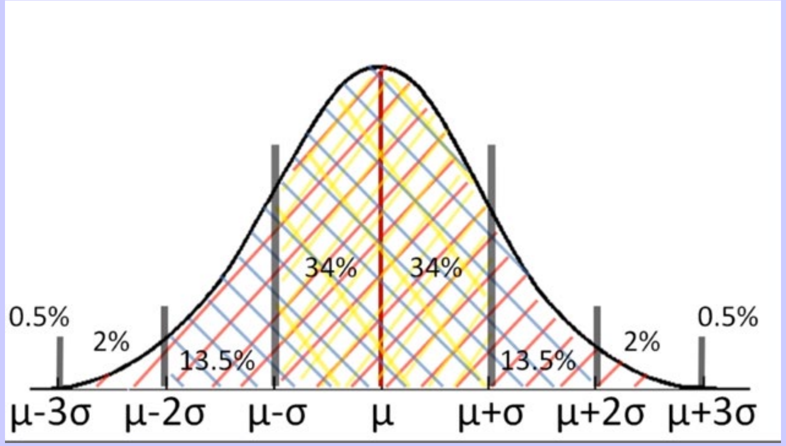
where μ and σ are the mean and standard deviation, respectively. The corresponding distribution function is given by $F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(v-\mu)^2/2\sigma^2} dv$

Example 12: **Assume that the current measurements in a strip of wire follows a normal distribution with a mean of 10 mA & a variance of 4 mA2. Let X denote the current in mA. What is the probability that a measurement exceeds 13 mA?**



Graphical probability that $X > 13$ for a normal random variable with $\mu = 10$ and $\sigma^2 = 4$.





A normal random variable with

$$\mu = 0 \ \ and \ \ \sigma^2 = 1$$

Is called a standard normal random variable and is denoted as Z.

$$Z = \frac{X - \mu}{\sigma}$$

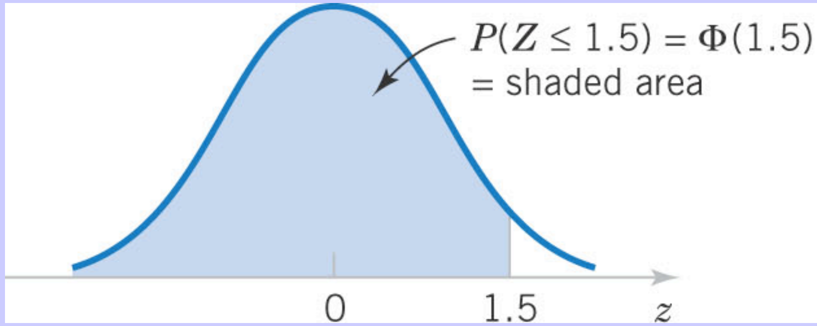
The cumulative distribution function of a standard normal random variable is denoted as:

$$\Phi(z) = P(Z \leq z) = F(z)$$

Values are found in Z Table.

Example 13: **Assume Z is a standard normal random variable. Find $P(Z \leq 1.50)$.**

Answer: 0.93319



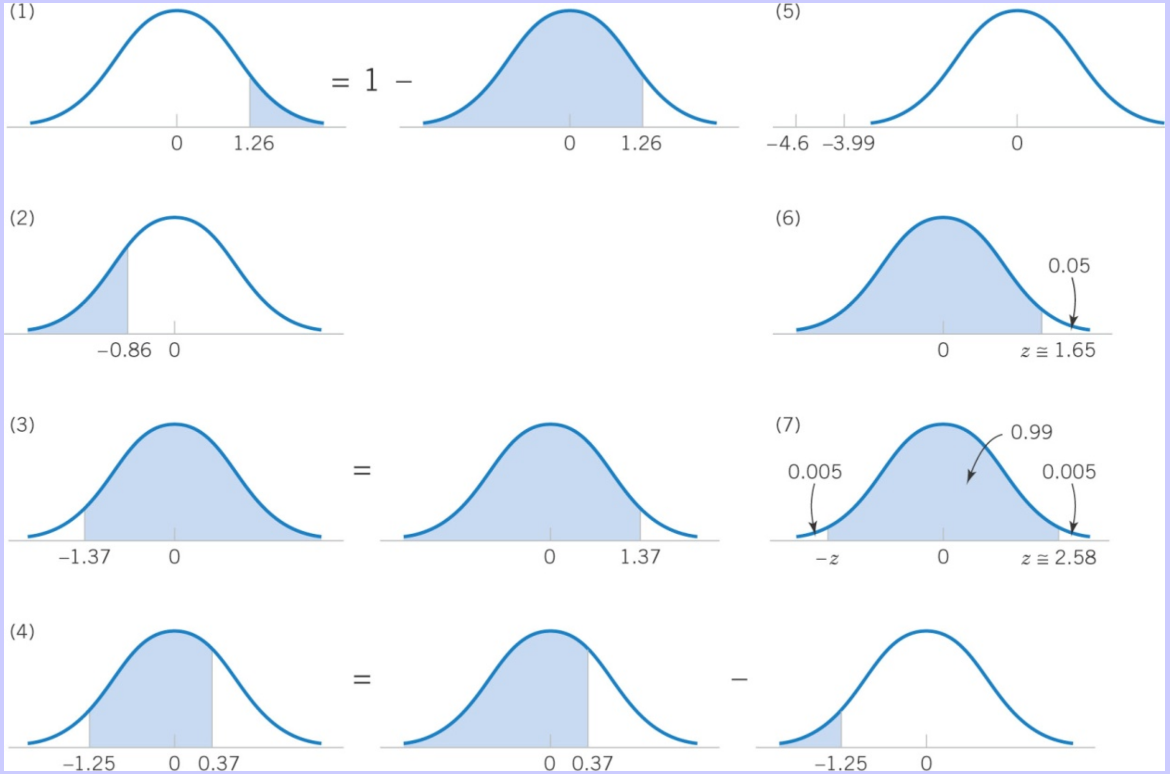
Standard normal PDF

z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Exercise 4: $P(Z \leq 1.53)$. **Answer: 0.93699**
Exercise 5: $P(Z \leq 0.01)$. **Answer: 0.50398**

Example 14:

1. P(Z > 1.26) = **0.1038**
2. P(Z < -0.86) = **0.195**
3. P(Z > -1.37) = **0.915**
4. P(-1.25 < 0.37) = **0.5387**
5. P(Z ≤ -4.6) ≈ **0**
6. Find z for P(Z ≤ z) = **0.05, z = -1.65**
7. Find z for (-z < Z < z) = **0.99, z = 2.58**

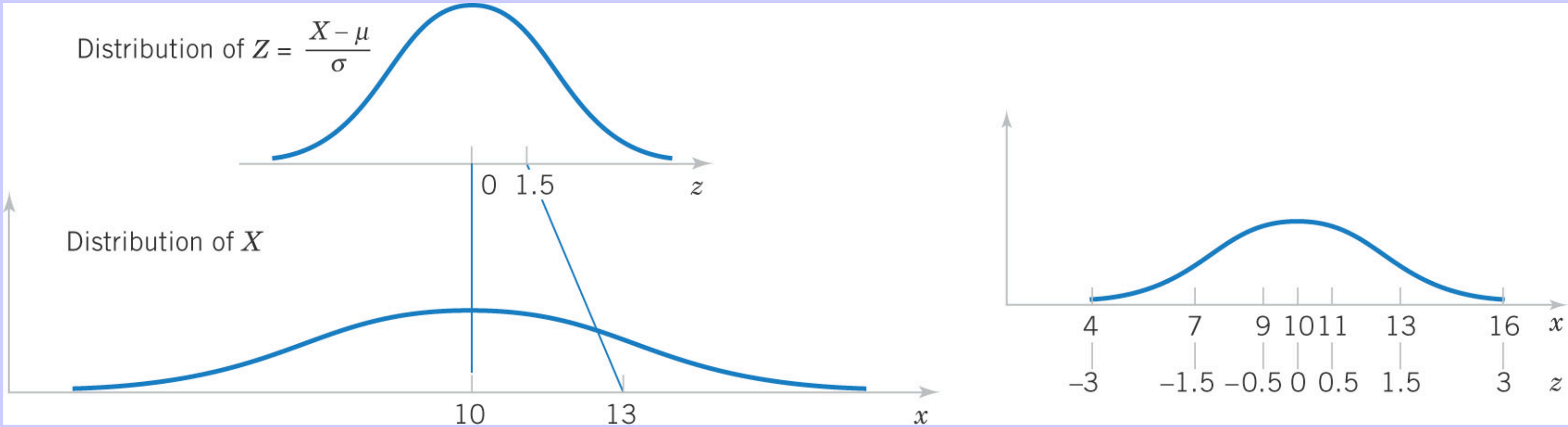


Graphical displays for standard normal distributions.

Example 15: **From a previous example with μ = 10 and σ = 2 mA, what is the probability that the current measurement is between 9 and 11 mA?**

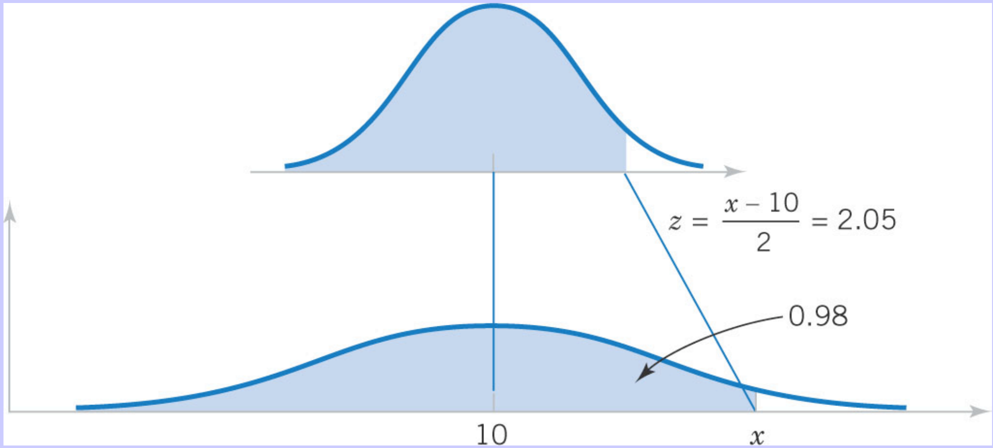
Answer:

$$P(9 < X < 11) = P\left(\frac{9-10}{2} < \frac{x-10}{2} < \frac{11-10}{2}\right)$$
$$= P(-0.5 < z < 0.5)$$
$$= P(z < 0.5) - P(z < -0.5)$$
$$= 0.69146 - 0.30854 = 0.38292$$



Example 16: Determine the value for which the probability that a current measurement is below this value is 0.98.

Answer:
$$P(X<x)=P\left(\frac{X-10}{2}<\frac{x-10}{2}\right)$$
$$=P\left(Z<\frac{x-10}{2}\right)=0.98$$
$z = 2.05$ is the closest value.
$$z=2(2.05)+10=14.1\text{ mA.}$$



Determining the value of x to meet a specified probability.

Homework 6: Assume that in the detection of a digital signal, the background noise follows a normal distribution with $\mu = 0$ volt and $\sigma = 0.45$ volt. The system assumes a digital 1 has been transmitted when the voltage exceeds 0.9.

1. What is the probability of detecting a digital 1 when none was sent? Let the random variable N denote the voltage of noise.
2. Determine the symmetric bounds about 0 that include 99% of all noise readings.
3. Suppose that when a digital 1 signal is transmitted, the mean of the noise distribution shifts to 1.8 volts. What is the probability that a digital 1 is not detected? Let S denote the voltage when a digital 1 is transmitted.

Homework 7: Repeat solve Homework 4 by using Gaussian distribution.

Homework 8: A television cable company receives numerous phone calls throughout the day from customers reporting service troubles and from would-be subscribers to the cable network. Most of these callers are put "on hold" until a company operator is free to help them. The company has determined that the length of time a caller is on hold is normally distributed with a mean of 3.1 minutes and a standard deviation 0.9 minutes. Company experts have decided that if as many as 5% of the callers are put on hold for 4.8 minutes or longer, more operators should be hired.

a) What proportion of the company's callers are put on hold for more than 4.8 minutes? Should the company hire more operators?
b) At another cable company (length of time a caller is on hold follows the same distribution as before), 2.5% of the callers are put on hold for longer than x minutes. Find the value of x

Homework 9: Suppose that a binary message either 0 or 1 must be transmitted by wire from location A to location B. However, the data sent over the wire are subject to a channel noise disturbance, so to reduce the possibility of error, the value 2 is sent over the wire when the message is 1 and the value - 2 is sent when the message is 0. If x, $x = \pm 2$, is the value sent from location A, then R, the value received at location B, is given by $R = x + N$, where N is the channel noise disturbance. When the message is received at location B the receiver decodes it according to the following rule:

If $R \geq 0.5$, then 1 is concluded

If $R < 0.5$, then 0 is concluded

If the channel noise follows the standard normal distribution compute the probability that the message will be wrong when decoded.

Let X be a discrete random variable that can take on the values 0, 1, 2,… such that the probability function of X is given by

$$f(x) = P(X = x) = \frac{\lambda^{x} e^{-\lambda}}{x\,!} \qquad x = 0, 1, 2, \ldots \qquad \lambda = np$$

where λ is a given positive constant. This distribution is called the Poisson distribution

Example 17: If the probability that an individual will suffer a bad reaction from injection of a given serum is 0.001, determine the probability that out of 2000 individuals, (a) exactly 3, (b) more than 2, individuals.

Solution:

X is Bernoulli distributed, but since bad reactions are assumed to be rare events, we can suppose that X is Poisson distributed

a) $P(X = x) = \dfrac{\lambda^{x} e^{-\lambda}}{x\,!}$ $\qquad \lambda = np = (2000)(0.001) = 2$

$P(X = 3) = \dfrac{\lambda^{3} e^{-2}}{3\,!} = 0.18$

b) $P(X>2) = 1 - [\,P(X=0)+P(X=1)+P(X=2)]$

$$= 1 - [\frac{2^0 e^{-2}}{0\,!} + \frac{2^1 e^{-2}}{1\,!} + \frac{2^2 e^{-2}}{2\,!}$$

$$= 1 - 5e^{-2} = 0.323$$

Exercise 6: Let's say you want to send a bit string of length $n = 10^4$ where each bit is independently corrupted with $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

Solution:

$\lambda = np = 10^4 10^{-6} = 0.01.$

$$P(X=0) = \frac{0.01^0 e^{-0.01}}{0\,!} = 0.99$$

Homework 10: Are we could have modelled X as a binomial distribution. That would have been computationally harder to compute but would have resulted in the same number (up to the millionth decimal)

Example 18: The number of visitors to a webserver per minute follows a Poisson distribution. If the average number of visitors per minute is 4, what is the probability that:

a) There are two or fewer visitors in one minute?

b) There are exactly two visitors in 30 seconds?

Solution:

a) The average number of visitors in a minute. In this case the parameter λ = 4. So the probability of two or fewer visitors in a minute is

$$P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X = 0) = \frac{e^{-4}4^0}{0!} = e^{-4}$$

$$P(X = 1) = \frac{e^{-4}4^1}{1!} = 4e^{-4}$$

$$P(X = 2) = \frac{e^{-4}4^2}{2!} = 8e^{-4}$$

The probability of two or fewer visitors in a minute is $e^{-4} + 4e^{-4} + 8e^{-4} = 0.238$

b) If the average number of visitors in 1 minute is 4, the average in 30 seconds is 2.

$$P(X = 2) = \frac{e^{-2}2^2}{2!} = 2e^{-2} = 0.271.$$

Homework 12: Consider a computer system with Poisson job-arrival stream at an average of 2 per minute. Determine the probability that in any one-minute interval there will be

a) 0 jobs;
b) exactly 2 jobs;
c) at most 3 arrivals.

Uniform Distribution

A random variable X is said to be uniformly distributed in $a \leq x \leq b$ if its density function is flat over a region.

$$f(x) = \frac{1}{b-a}, \qquad for\ a \leq x \leq b$$

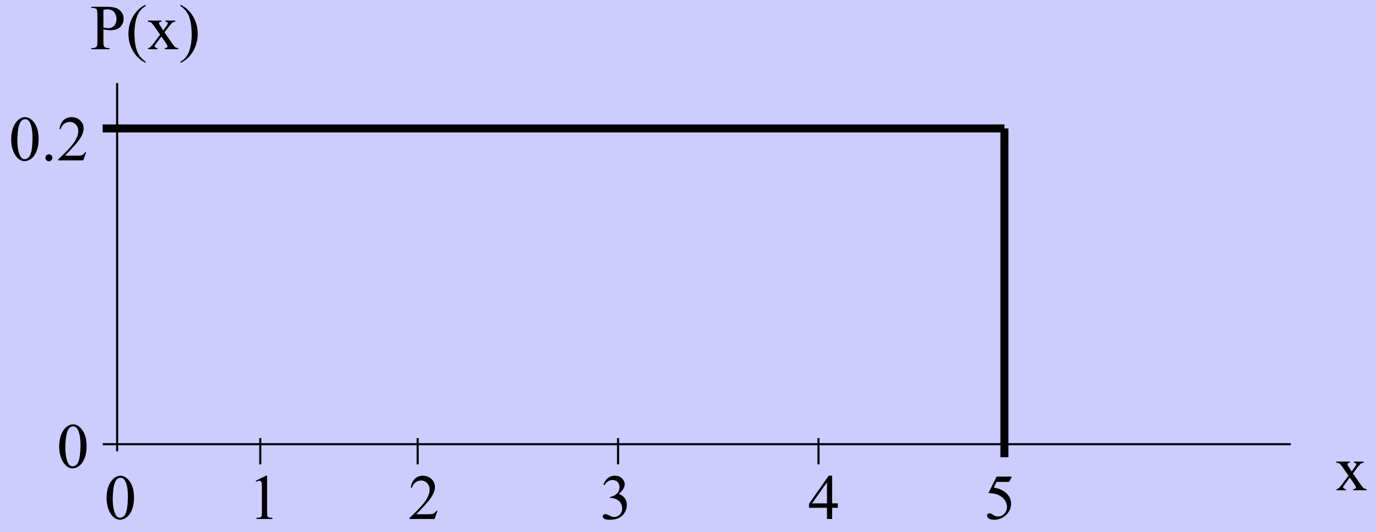
The distribution function is given by

$$F(x)=P(X \leq x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a), & a \leq x \leq b \\ 1 & x \geq b \end{cases}$$

The mean and variance are, respectively

$$\mu_x = \frac{1}{2}(a+b), \qquad \sigma^2 = \frac{1}{12}(b-a)^2$$

Example 19: The figure below depicts the probability distribution for temperatures in a manufacturing process. The temperatures are controlled so that they range between 0 and 5 degrees Celsius, and every possible temperature is equally likely.



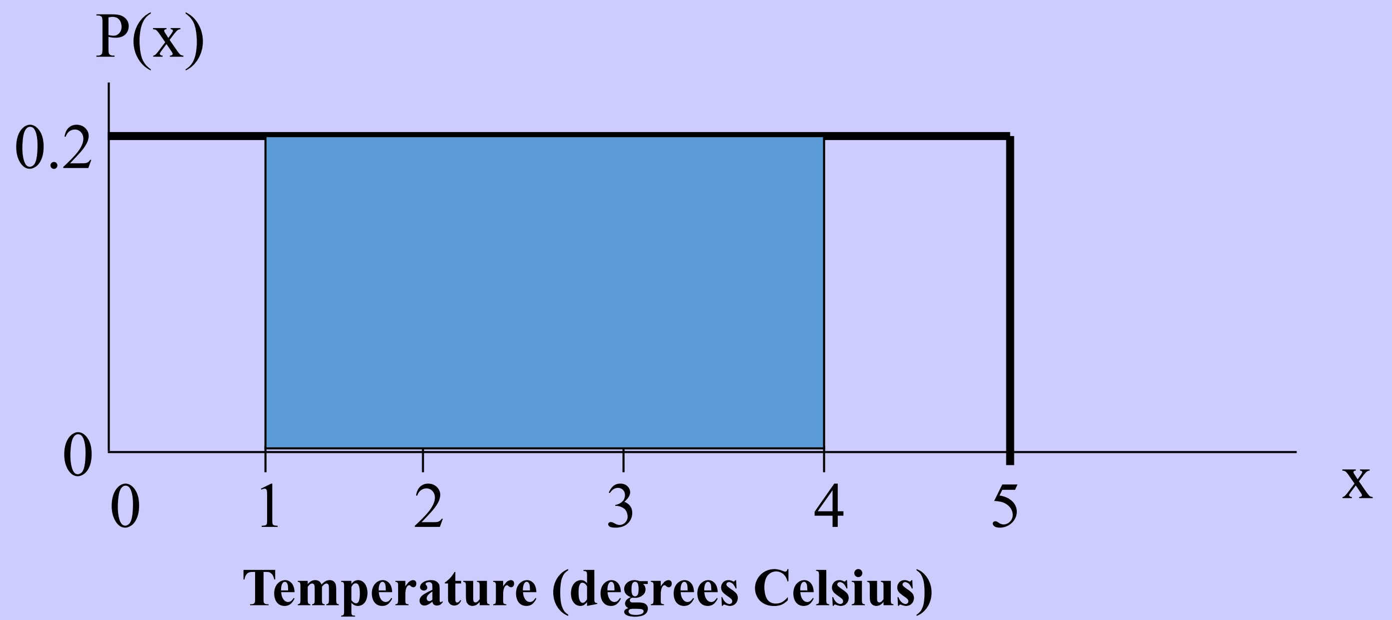
Temperature (degrees Celsius)

1. **What is the Probability that the temperature is exactly 4 degrees?** **Answer: 0**

   Since we have a continuous random variable there are an infinite number of possible outcomes between 0 and 5, the probability of one number out of an infinite set of numbers is 0.

2. What is the probability the temperature is between 1C and 4C?

Answer:



The total area of the rectangle is 1, and we can see that the part of the rectangle between 1 and 4 is 3/5 of the total, so $P(1 \le x \le 4) = 3/5*(1) = 0.6$.

Homework 13: **The current (in mA) measured in a piece of copper wire is known to follow a uniform distribution over the interval [0, 25]. Write down the formula for the probability density function f(x) of the random variable X representing the current. Calculate the mean and variance of the distribution and find the cumulative distribution function F(x).**