

ECS315 2019/1 Part IV.1 Dr.Prapun

10 Continuous Random Variables

10.1 From Discrete to Continuous Random Variables

10.1. In many practical applications of probability, physical situations are better described by random variables that can take on a *continuum* of possible values rather than a *discrete* number of values.

For the random variables to be discussed in this section,

- any individual value has probability zero:

$$P[X = x] = 0 \quad \text{for all } x \quad (20)$$

and that

- the supports are always uncountable.

These random variables are called **continuous random variables**.

10.2. Implications:

- (a) We can see from (20) that the pmf is going to be useless for this type of random variable. It turns out that the cdf F_X is still useful and we shall introduce another useful function called *probability density function* (pdf) to replace the role of pmf. However, integral calculus³⁷ is required to formulate this continuous analog of a pmf.

³⁷This is always a difficult concept for the beginning student.



- (b) Because talking about $P[X = x]$ for continuous RV is useless (always 0), we instead talk about the probability that the RV is in some interval, e.g. $P[a < X < b]$.

10.3. In some cases, the random variable X is actually discrete but, because the range of possible values is so large, it might be more convenient to analyze X as a continuous random variable.

Example 10.4. Suppose that current measurements are read from a digital instrument that displays the current to the nearest one-hundredth of a mA. Because the possible measurements are limited, the random variable is discrete. However, it might be a more convenient, simple approximation to assume that the current measurements are values of a continuous random variable.

Example 10.5. If you can measure the heights of people with infinite precision, the height of a randomly chosen person is a continuous random variable. In reality, heights cannot be measured with infinite precision, but the mathematical analysis of the distribution of heights of people is greatly simplified when using a mathematical model in which the height of a randomly chosen person is modeled as a continuous random variable. [21, p 284]

Example 10.6. Continuous random variables are important models for

- (a) voltages in communication receivers
- (b) file download times on the Internet
- (c) velocity and position of an airliner on radar
- (d) lifetime of a battery
- (e) decay time of a radioactive particle
- (f) time until the occurrence of the next earthquake in a certain region
- (g) noise in communication systems

Example 10.7. The simplest example of a continuous random variable is the “random choice” of a number from the interval $(0, 1)$.

- In MATLAB, this can be generated by the command `rand`. In Excel, use `rand()`.
- The generation is “unbiased” in the sense that “any number in the range $(0,1)$ is as likely to occur as another number.”
- Histogram is flat over $(0,1)$ in the limit as the number of samples increases to infinity regardless of the number of bins as long as the bins have the same size. See Figure 25b.
- Formally, this is called a uniform RV on the interval $(0, 1)$.

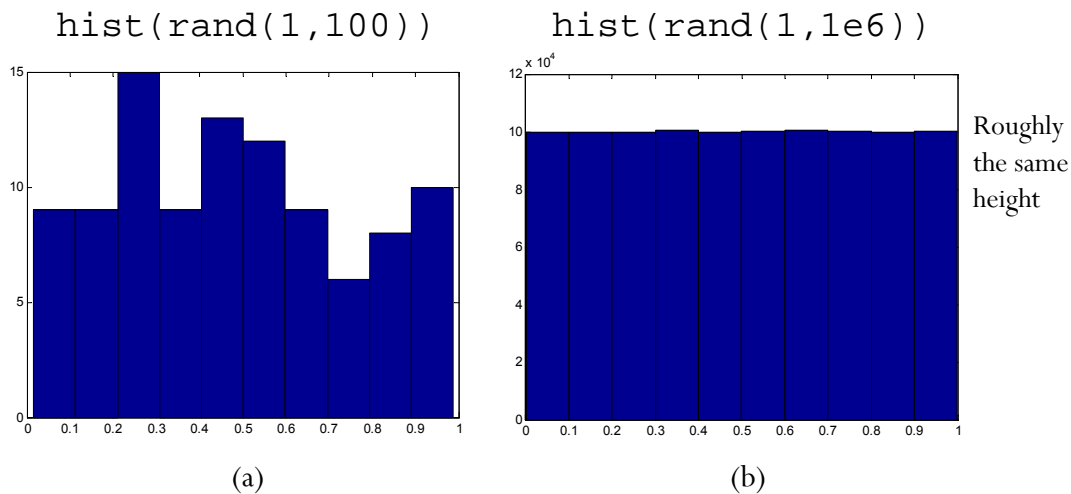


Figure 25: Histogram of the values generated by MATLAB command `rand`.

Example 10.8. Put a piece of (unit-width and unit-height) paper outdoor. Mark the location of the (center of) first drop of rain on it. Record its horizontal position).

Example 10.9. In MATLAB, there are other commands (such as `randn`) and ways to generate continuous random variables with other shapes of histograms.

Definition 10.10. We say that X is a **continuous random variable**³⁸ if we can find a (real-valued) function³⁹ f such that, for any set B , $P[X \in B]$ has the form

$$P[X \in B] = \int_B f(x)dx. \quad (21)$$

Equivalently,

$$P[\text{some statement(s) about } X] = \int_{\substack{\{\text{all the } x \text{ values that} \\ \text{satisfy the statement(s)}\}} f(x) dx.$$

- In particular,

$$P[a \leq X \leq b] = \int_a^b f(x)dx. \quad (22)$$

In other words, the **area under the graph** of $f(x)$ between the points a and b gives the probability $P[a \leq X \leq b]$.

- The function f is called the **probability density function** (pdf) or simply **density**.
- When we want to emphasize that the function f is a density of a particular random variable X , we write f_X instead of f .

³⁸To be more rigorous, this is the definition for *absolutely* continuous random variable. At this level, we will not distinguish between the continuous random variable and absolutely continuous random variable. When the distinction between them is considered, a random variable X is said to be continuous (not necessarily absolutely continuous) when condition (20) is satisfied. Alternatively, condition (20) is equivalent to requiring the cdf F_X to be continuous. Another fact worth mentioning is that if a random variable is absolutely continuous, then it is continuous. So, absolute continuity is a stronger condition.

³⁹Strictly speaking, δ -“function” is not a function; so, can’t use δ -function here.

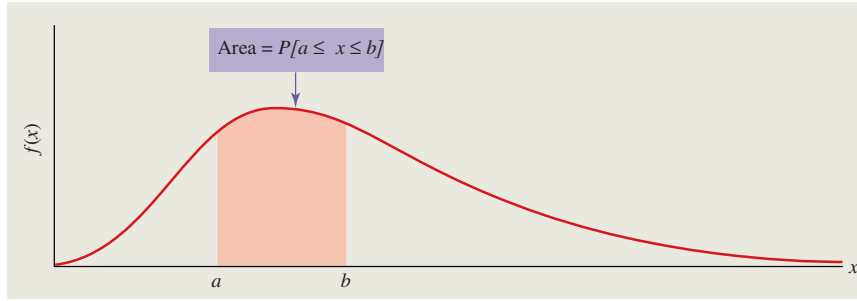


Figure 26: For a continuous random variable, the probability distribution is described by a curve called the probability density function, $f(x)$. The total area beneath the curve is 1.0, and the probability that X will take on some value between a and b is the area beneath the curve between points a and b .

Example 10.11. For the random variable generated by the `rand` command in MATLAB⁴⁰ or the `rand()` command in Excel,

Definition 10.12. Recall that the support S_X of a random variable X is any set S such that $P[X \in S] = 1$. For continuous random variable, S_X is usually set to be $\{x : f_X(x) > 0\}$.

Example 10.13. For the random variable X in Example 10.11,

Example 10.14. For noise in communication systems,

⁴⁰The `rand` command in MATLAB is an approximation for two reasons:

- (a) It produces pseudorandom numbers; the numbers seem random but are actually the output of a deterministic algorithm.
- (b) It produces a double precision floating point number, represented in the computer by 64 bits. Thus MATLAB distinguishes no more than 2^{64} unique double precision floating point numbers. By comparison, there are uncountably infinite real numbers in the interval from 0 to 1.

Example 10.15. Consider a random variable X whose pdf is

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

- (a) Find $P[X > 0.5]$.

- (b) Find $P[0.2 < X < 0.3]$.

- (c) Find $P[0.19 < X < 0.21]$.

- (d) Find $P[0.79 < X < 0.81]$.

Observation: From the pdf expression, we know that $f_X(0.8) > f_X(0.2)$.

- (a) Does this imply $P[X = 0.8] > P[X = 0.2]$? No! From (20), we know that both probabilities are 0.

- (b) $f_X(0.8) > f_X(0.2)$ simply means the RV X is more likely to be in the small interval around 0.8 than in the small interval (of the same length) around 0.2. In fact, the ratio of the two probabilities is approximately the ratio of the pdf values.

10.16. Intuition/Interpretation:

The use of the word “density” originated with the analogy to the distribution of matter in space. In physics, any finite volume, no matter how small, has a positive mass, but there is no mass at a single point. A similar description applies to continuous random variables.

Approximately, for a small Δx ,

$$P[X \in [x, x + \Delta x]] = \int_x^{x+\Delta x} f_X(t) dt \approx f_X(x) \Delta x.$$

This is why we call f_X the density function.

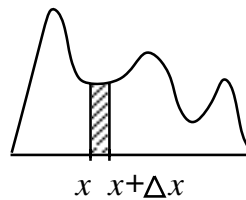


Figure 27: $P[x \leq X \leq x + \Delta x]$ is the area of the shaded vertical strip.

In other words, the probability of random variable X taking on a value in a *small* interval around point c is approximately equal to $f(c) \times d$ when d is the length of the interval.

- In fact, $f_X(x) = \lim_{\Delta x \rightarrow 0} \frac{P[x < X \leq x + \Delta x]}{\Delta x}$
- The number $f_X(x)$ itself is **not a probability**. In particular, it does not have to be between 0 and 1.
- $f_X(c)$ is a relative measure for the likelihood that random variable X will take on a value in the immediate neighborhood of point c .

Stated differently, the pdf $f_X(x)$ expresses how densely the probability mass of random variable X is smeared out in the neighborhood of point x . Hence, the name of density function.

10.17. Histogram and pdf [21, p 143 and 145]:

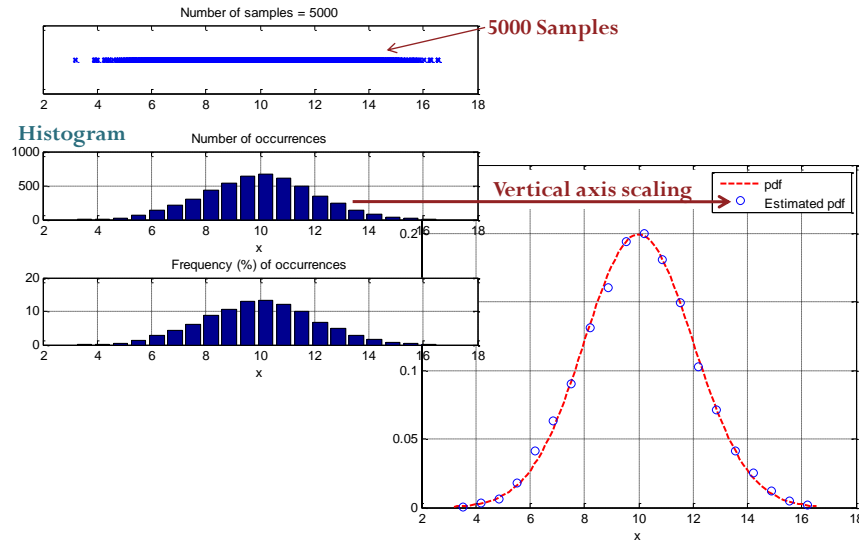


Figure 28: From histogram to pdf.

- (a) A probability **histogram** is a bar chart that divides the range of values covered by the samples/measurements into intervals of the same width, and shows the proportion (relative frequency) of the samples in each interval.
- To make a histogram, break up the range of values covered by the samples into a number of disjoint adjacent intervals each having the same width, say width Δ . The height of the bar on each interval $[j\Delta, (j + 1)\Delta)$ is taken such that the area of the bar is equal to the proportion of the measurements falling in that interval (the proportion of measurements within the interval is divided by the width of the interval to obtain the height of the bar).
 - The total area under the probability histogram is thus standardized/normalized to one.
- (b) If you take sufficiently many independent samples from a continuous random variable and make the width Δ of the base intervals of the probability histogram smaller and smaller, the graph of the probability histogram will begin to look more and more like the pdf.
- (c) Conclusion: A probability density function can be seen as a “smoothed-out” normalized version of a (probability) histogram

10.2 Properties of PDF and CDF for Continuous Random Variables

10.18. $p_X(x) = P[X = x] = P[x \leq X \leq x] = \int_x^x f_X(t)dt = 0.$

Again, it makes no sense to speak of the probability that X will take on a pre-specified value. This probability is always zero.

10.19. $P[X = a] = P[X = b] = 0.$ Hence,

$$P[a < X < b] = P[a \leq X < b] = P[a < X \leq b] = P[a \leq X \leq b]$$

- The corresponding integrals over an interval are not affected by whether or not the endpoints are included or excluded.⁴¹

10.20. The pdf f_X is determined only almost everywhere⁴². Given a pdf f for a continuous random variable X , if we construct a function g by changing the function f at a countable number of points⁴³, then g can also serve as a pdf for X .

This is because f_X is defined via its integration property. Changing the value of a function at a few points does not change its area under the curve (from a to b)

10.21. The cdf of any kind of random variable X is defined as

$$F_X(x) = P[X \leq x].$$

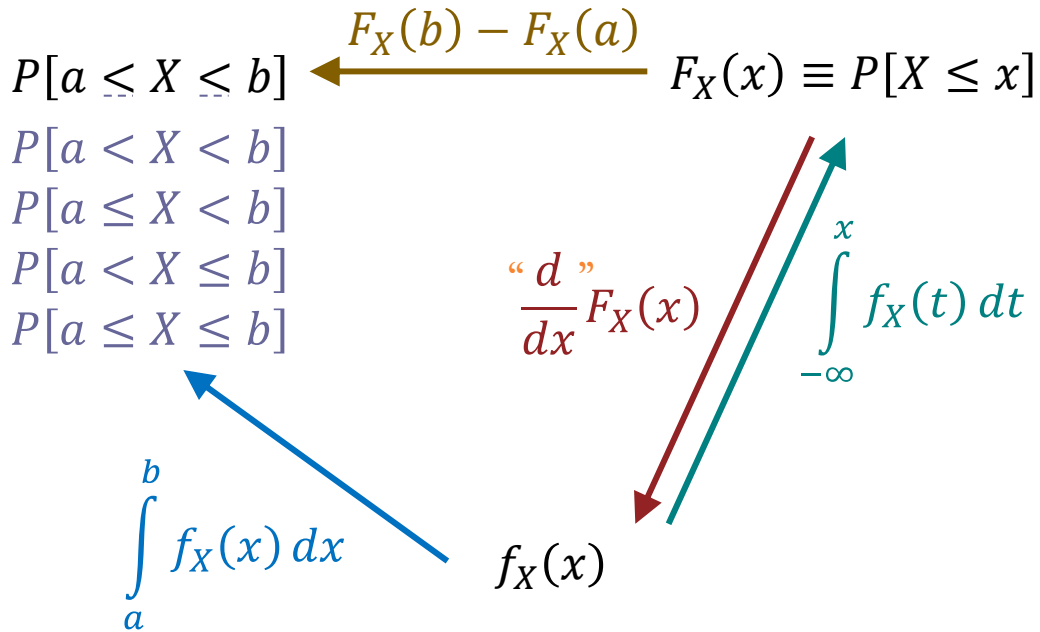
Note that even though there are more than one valid pdfs for any given random variable, the cdf is unique. There is only one cdf for each random variable.

⁴¹This implies that, when we work with continuous random variables, it is usually not necessary to be precise about specifying whether or not a range of numbers includes the endpoints. This is quite different from the situation we encounter with discrete random variables where it is critical to carefully examine the type of inequality.

⁴²Lebesgue-a.e., to be exact

⁴³More specifically, if $g = f$ Lebesgue-a.e., then g is also a pdf for X .

(a) pdf, cdf, and probability calculation for continuous RV



(b) Finding Probabilities from CDF

Definition: $F_X(x) \equiv P[X \leq x]$

For any RV ,	For continuous RV ,
<ul style="list-style-type: none"> $P[X \leq b] = F_X(b)$ $P[X < b] = F_X(b) - P[X = b]$ 	<ul style="list-style-type: none"> $P[X \leq b] = F_X(b)$ $P[X < b] = F_X(b)$
<ul style="list-style-type: none"> $P[X > a] = 1 - F_X(a)$ $P[X \geq a] = 1 - F_X(a) + P[X = a]$ 	<ul style="list-style-type: none"> $P[X > a] = 1 - F_X(a)$ $P[X \geq a] = 1 - F_X(a)$
<ul style="list-style-type: none"> $P[a < X \leq b] = F_X(b) - F_X(a)$ 	<ul style="list-style-type: none"> $P[a < X \leq b] = F_X(b) - F_X(a)$ $P[a < X < b] = F_X(b) - F_X(a)$ $P[a \leq X < b] = F_X(b) - F_X(a)$ $P[a \leq X \leq b] = F_X(b) - F_X(a)$
<ul style="list-style-type: none"> $P[X = a] = F_X(a) - F_X(a^-)$ (amount of jump in the CDF @ a) 	<ul style="list-style-type: none"> $P[X = a] = 0$

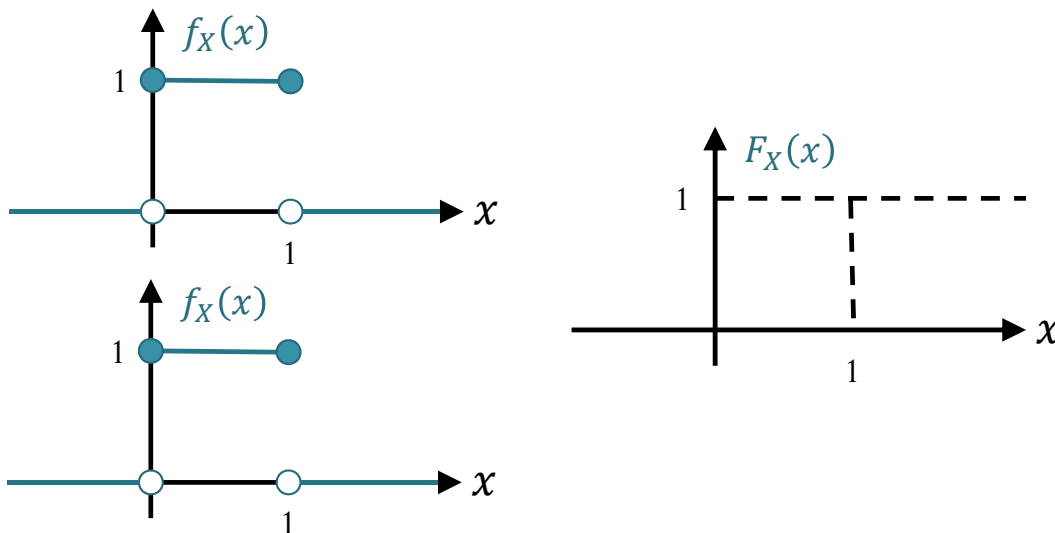
Figure 29: Summary of properties involving CDF

10.22. Unlike the cdf of a discrete random variable, the cdf of a continuous random variable has no jump and is continuous everywhere.

10.23. For continuous random variable, given the pdf $f_X(x)$, we can find the cdf of X by

$$F_X(x) = P[X \leq x] = \int_{-\infty}^x f_X(t) dt.$$

Example 10.24. For the random variable generated by the rand command in MATLAB or the rand() command in Excel,



10.25. Given the cdf $F_X(x)$ of a continuous random variable, we can find the pdf $f_X(x)$ by

Step 1 If F_X is differentiable at x , we set

$$\frac{d}{dx} F_X(x) = f_X(x).$$

Step 2 From 10.20, at countably many points, we can set the values of f_X to be any value. We use this to deal with the boundary

point(s) including the point(s) where F_X is not differentiable. Usually, the values are selected to give simple expression. (In many cases, they are simply set to 0.)

Example 10.26. Suppose that the lifetime X of a device has the cdf

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{4}x^2, & 0 \leq x \leq 2, \\ 1, & x > 2. \end{cases}$$

Because $F_x(x)$ is defined piecewise and the expression defining each piece is “nice”, we can find the derivative for each piece and get

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{2}, & 0 < x < 2, \\ 0, & x > 2. \end{cases}$$

This leaves two points⁴⁴ to be considered: $x = 0$ and $x = 2$. However, they are only two points and therefore, from 10.20, the values of the pdf can be any real numbers. Here, we set the values to be 0 at both points:

$$f_X(x) = \begin{cases} \frac{1}{2}x, & 0 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

10.27. In many situations when you are asked to find the pdf from a description of a random variable, it may be easier to find cdf first and then differentiate it to get pdf.

Exercise 10.28. A point is “picked at random” in the inside of a circular disk with radius r . Let the random variable X denote the distance from the center of the disk to this point. Find $f_X(x)$.

⁴⁴At each of these boundary points, the expressions on both of its sides are different and hence, to really find its derivative, we need to consider whether the derivative from the left exists and is the same as the derivative from the right. At $x = 0$, turn out that the slope on both sides is 0. So the derivative exists. However, at $x = 2$, F_X has no derivative: the slope is 1 from the left but 0 from the right.

10.29. f_X is nonnegative and $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Example 10.30. Random variable X has pdf

$$f_X(x) = \begin{cases} ce^{-2x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the constant c and sketch the pdf.

Example 10.31. The pdf of a random variable X is shown in Figure 30. What should be the value of h ?

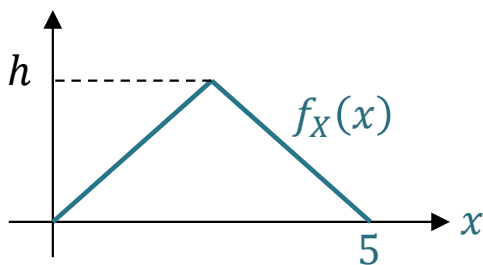


Figure 30: Triangular pdf for Example 10.31.

Definition 10.32. A continuous random variable is called *exponential* if its pdf is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0 \end{cases}$$

for some $\lambda > 0$

Example 10.33. In Example 10.30, X is an exponential random variable with $\lambda = 2$.

Theorem 10.34. Any nonnegative⁴⁵ function that integrates to one is a *probability density function* (pdf) of some random variable [9, p.139].

⁴⁵or nonnegative a.e.

10.3 Expectation and Variance

10.35. *Expectation:* Suppose X is a continuous random variable with probability density function $f_X(x)$.

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x) dx \quad (23)$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (24)$$

In particular,

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx$$
$$\text{Var } X = \int_{-\infty}^{\infty} (x - \mathbb{E}X)^2 f_X(x) dx = \mathbb{E}[X^2] - (\mathbb{E}X)^2.$$

Example 10.36. For the random variable generated by the `rand` command in MATLAB or the `rand()` command in Excel,

Example 10.37. For the exponential random variable introduced in Definition 10.32,

10.38. If we compare other characteristics of discrete and continuous random variables, we find that with discrete random variables, many facts are expressed as sums. With continuous random variables, the corresponding facts are expressed as integrals.

10.39. All of the properties for the expectation and variance of discrete random variables also work for continuous random variables as well:

- (a) Intuition/interpretation of the expected value: As $n \rightarrow \infty$, the average of n independent samples of X will approach $\mathbb{E}X$. This observation is known as the “Law of Large Numbers”.
- (b) For $c \in \mathbb{R}$, $\mathbb{E}[c] = c$
- (c) For constants a, b , we have $\mathbb{E}[aX + b] = a\mathbb{E}X + b$.
- (d) $\mathbb{E}[\sum_{i=1}^n c_i g_i(X)] = \sum_{i=1}^n c_i \mathbb{E}[g_i(X)]$.
- (e) $\text{Var } X = \mathbb{E}[X^2] - (\mathbb{E}X)^2$
- (f) $\text{Var } X \geq 0$.
- (g) $\text{Var } X \leq \mathbb{E}[X^2]$.
- (h) $\text{Var}[aX + b] = a^2 \text{Var } X$.
- (i) $\sigma_{aX+b} = |a| \sigma_X$.

10.40. Chebyshev’s Inequality:

$$P[|X - \mathbb{E}X| \geq \alpha] \leq \frac{\sigma_X^2}{\alpha^2}$$

or equivalently

$$P[|X - \mathbb{E}X| \geq n\sigma_X] \leq \frac{1}{n^2}$$

- This inequality use variance to bound the “tail probability” of a random variable.
- Useful only when $\alpha > \sigma_X$

Example 10.41. A circuit is designed to handle a current of 20 mA plus or minus a deviation of less than 5 mA. If the applied current has mean 20 mA and variance 4 (mA)^2 , use the Chebyshev inequality to bound the probability that the applied current violates the design parameters.

Let X denote the applied current. Then X is within the design parameters if and only if $|X - 20| < 5$. To bound the probability that this does not happen, write

$$P[|X - 20| \geq 5] \leq \frac{\text{Var } X}{5^2} = \frac{4}{25} = 0.16.$$

Hence, the probability of violating the design parameters is at most 16%.

10.42. Interesting applications of expectation:

(a) $f_X(x) = \mathbb{E}[\delta(X - x)]$

(b) $P[X \in B] = \mathbb{E}[1_B(X)]$

Example 10.43. Consider two distributions for a random variable X . In part (a), which corresponds to the second column in the table below, X is a *discrete* random variable with its pmf specified in the first row. In part (b), which corresponds to the third column, X is a *continuous* random variable with its pdf specified in the first row.

Distribution	$p_X(x) = \begin{cases} cx^2, & x \in \{1, 2\}, \\ 0, & \text{otherwise.} \end{cases}$	$f_X(x) = \begin{cases} cx^2, & x \in (1, 2), \\ 0, & \text{otherwise.} \end{cases}$
(i) Find c		
(ii) Find $\mathbb{E}X$		
(iii) Find $\mathbb{E}[X^2]$		
(iv) Find $\text{Var } X$		
(v) Find σ_X		

ECS315 2019/1 Part IV.2 Dr.Prapun

10.4 Families of Continuous Random Variables

Theorem 10.34 states that any nonnegative function $f(x)$ whose integral over the interval $(-\infty, +\infty)$ equals 1 can be regarded as a probability density function of a random variable. In real-world applications, however, special mathematical forms naturally show up. In this section, we introduce a couple families of continuous random variables that frequently appear in practical applications. The probability densities of the members of each family all have the same mathematical form but differ only in one or more parameters.

10.4.1 (Continuous) Uniform Distribution

Definition 10.44.

- The (continuous) uniform distribution on an interval $[a, b]$, is denoted by $\text{uniform}([a, b])$ or $\mathcal{U}([a, b])$ or simply $\mathcal{U}(a, b)$.
- Expressions that are synonymous with “ X is a uniform random variable” are
 - (a) “ X is uniformly distributed”,
 - (b) “ X has a uniform distribution”,
 - (c) and “ X has a uniform density”.

To specify the support (range) of X , we may also append “on/over the interval (a, b) ”.

- This family is characterized by pdf of the form

$$f_X(x) = \begin{cases} 0, & x < a, x > b \\ \frac{1}{b-a}, & a \leq x \leq b \end{cases}$$

- The constants a and b are referred to as the parameters of the uniform distribution.

10.45. Important Interpretation: A continuous uniform random variable X on the interval $[a, b]$ is just as likely to be “near” any value in $[a, b]$ as any other value.

10.46. In MATLAB,

- use $X = a+(b-a)*\text{rand}$ or $X = \text{random}(\text{'Uniform'}, a, b)$ to generate $X \sim \mathcal{U}(a, b)$,
- use $\text{pdf}(\text{'Uniform'}, x, a, b)$ and $\text{cdf}(\text{'Uniform'}, x, a, b)$ to evaluate the pdf and cdf at x , respectively.

Exercise 10.47. Show that $F_X(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$

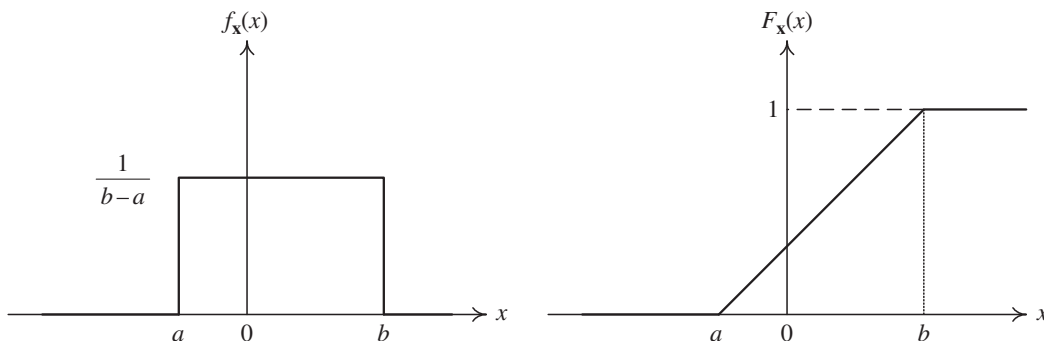


Figure 31: The pdf and cdf for the uniform random variable. [16, Fig. 3.5]

Example 10.48 (F2011). Suppose X is uniformly distributed on the interval $(1, 2)$. ($X \sim \mathcal{U}(1, 2)$.)

(a) Plot the pdf $f_X(x)$ of X .

(b) Plot the cdf $F_X(x)$ of X .

10.49. The uniform distribution provides a probability model for selecting a point at random from the interval $[a, b]$.

- Use with caution to model a quantity that is known to vary randomly between a and b but about which little else is known.

Example 10.50. [9, Ex. 4.1 p. 140-141] In coherent radio communications, the phase difference between the transmitter and the receiver, denoted by Θ , is modeled as having a uniform density on $[-\pi, \pi]$.

(a) $P[\Theta \leq 0] = \frac{1}{2}$

(b) $P[\Theta \leq \frac{\pi}{2}] = \frac{3}{4}$

Exercise 10.51. Show that when $X \sim \mathcal{U}(a, b)$, $\mathbb{E}X = \frac{a+b}{2}$, $\text{Var } X = \frac{(b-a)^2}{12}$, and $\mathbb{E}[X^2] = \frac{1}{3}(b^2 + ab + a^2)$.

10.4.2 Gaussian Distribution

10.52. This is the most widely used model for the distribution of a random variable. When you have many independent random variables, a fundamental result called the central limit theorem (CLT) (informally) says that the sum (or the average) of them can often be approximated by normal distribution.

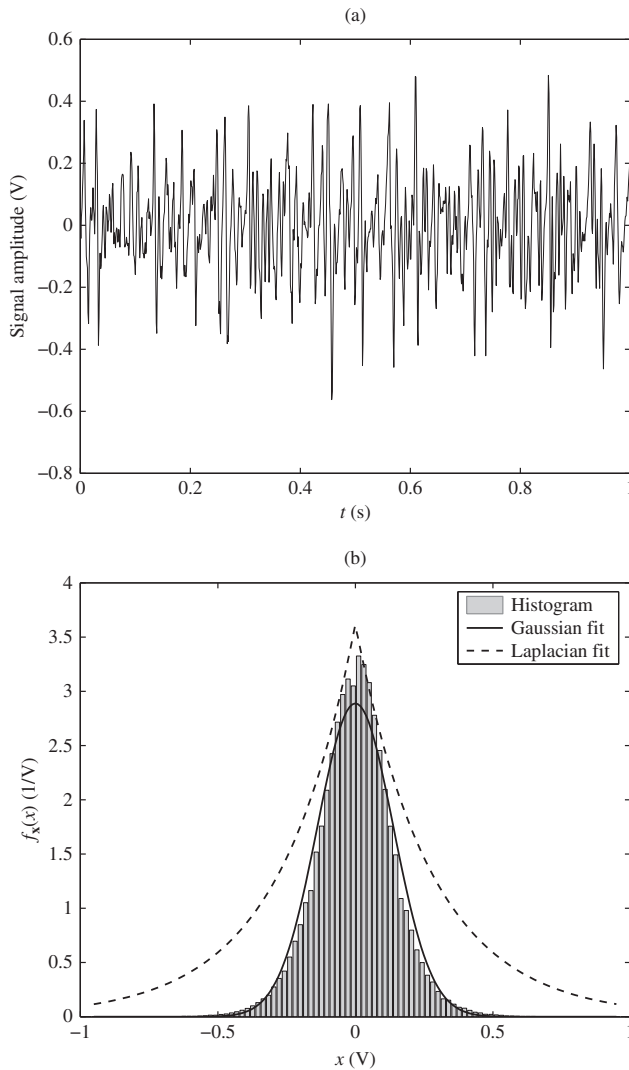


Figure 32: Electrical activity of a skeletal muscle: (a) A sample skeletal muscle (emg) signal, and (b) its histogram and pdf fits. [16, Fig. 3.14]

Definition 10.53. *Gaussian* random variables:

- Often called *normal* random variables because they occur so frequently in practice.

- The Gaussian distribution is denoted by $\mathcal{N}(m, \sigma^2)$. It has two parameters: $m \in \mathbb{R}$ and $\sigma > 0$.
 - Caution: The second argument in $\mathcal{N}(m, \sigma^2)$ is σ^2 (not σ).
 - Several references use μ instead of m .
- This family is characterized by pdf of the form

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}.$$

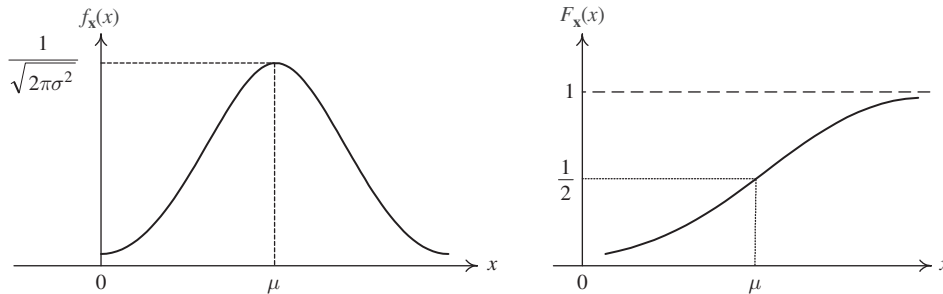


Figure 33: The pdf and cdf of $\mathcal{N}(\mu, \sigma^2)$. [16, Fig. 3.6]

- In Excel, use `NORMDIST(x,m,σ, FALSE)`.
In MATLAB, use `normpdf(x,m,σ)` or `pdf('Normal',x,m,σ)`.
- Figure 33 and Figure 35 display the famous **bell-shaped** graphs of the Gaussian pdf. These curves are also called the *normal* curves.
- In MATLAB, use `X = random('Normal',m,σ)` or `X = σ*randn + m` to generate $X \sim \mathcal{N}(m, \sigma^2)$.
- $F_X(x)$ has no closed-form expression. However, see 10.60.
 - In MATLAB, use `normcdf(x,m,σ)` or `cdf('Normal',x,m,σ)`.
 - In Excel, use `NORMDIST(x,m,σ, TRUE)`.

10.54. $\mathbb{E}X = m$ and $\text{Var } X = \sigma^2$.

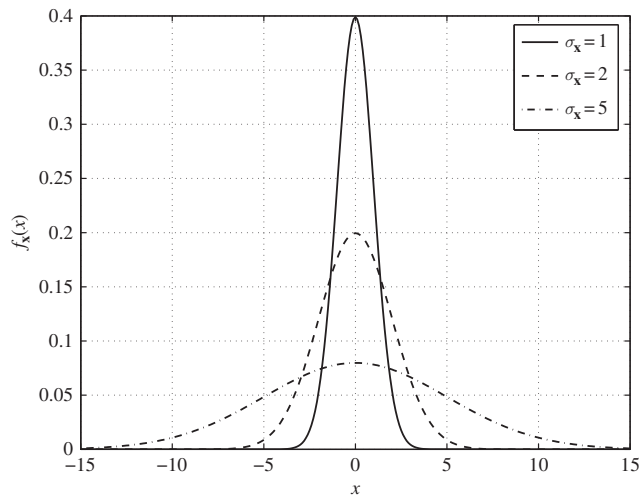


Figure 34: Plots of the zero-mean Gaussian pdf for different values of standard deviation, σ_X . [16, Fig. 3.15]

10.55. Important probabilities:

$$P[|X - \mu| < \sigma] = 0.6827;$$

$$P[|X - \mu| > \sigma] = 0.3173;$$

$$P[|X - \mu| > 2\sigma] = 0.0455;$$

$$P[|X - \mu| < 2\sigma] = 0.9545$$

These values are illustrated in Figure 35. We will see these numbers again in Example 10.61.

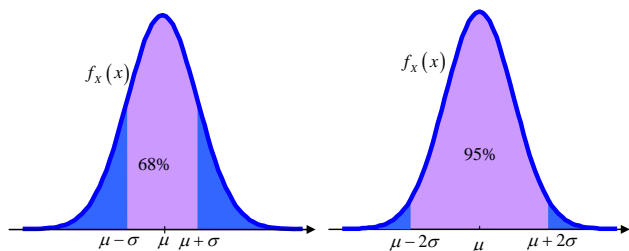


Figure 35: Probability density function of $X \sim \mathcal{N}(\mu, \sigma^2)$. The purple areas correspond to $P[|X - \mu| < \sigma] = 0.6827$ and $P[|X - \mu| < 2\sigma] = 0.9545$, respectively.

Example 10.56. Figure 36 compares several deviation scores and the normal distribution.

- (a) Standard scores have a mean of zero and a standard deviation of 1.0.
- (b) Scholastic Aptitude Test scores have a mean of 500 and a standard deviation of 100.

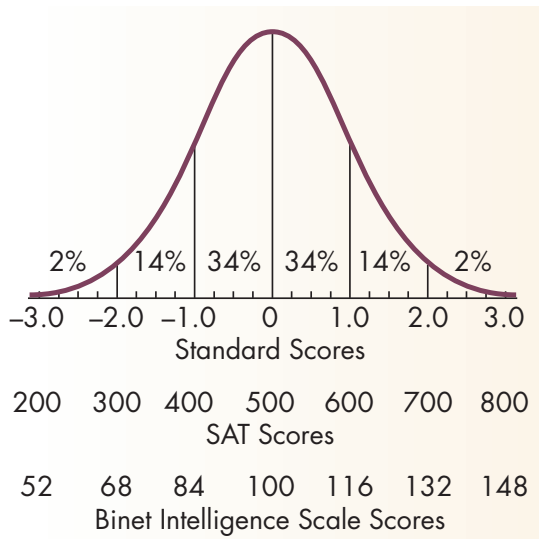


Figure 36: Comparison of Several Deviation Scores and the Normal Distribution

(c) Binet Intelligence Scale⁴⁶ scores have a mean of 100 and a standard deviation of 16.

In each case there are 34 percent of the scores between the mean and one standard deviation, 14 percent between one and two standard deviations, and 2 percent beyond two standard deviations. [Source: Beck, *Applying Psychology: Critical and Creative Thinking*.]

10.57. The area under a normal probability density function beyond 3σ from the mean is quite small. In fact,

$$P[|X - \mu| < 3\sigma] \approx 0.9973.$$

Therefore, approximately 99.73% of the probability of a normal distribution is within the interval $(\mu - 3\sigma, \mu + 3\sigma)$.

⁴⁶Alfred Binet, who devised the first general aptitude test at the beginning of the 20th century, defined intelligence as the ability to make adaptations. The general purpose of the test was to determine which children in Paris could benefit from school. Binet's test, like its subsequent revisions, consists of a series of progressively more difficult tasks that children of different ages can successfully complete. A child who can solve problems typically solved by children at a particular age level is said to have that mental age. For example, if a child can successfully do the same tasks that an average 8-year-old can do, he or she is said to have a mental age of 8. The intelligence quotient, or IQ, is defined by the formula:

$$IQ = 100 \times (\text{Mental Age} / \text{Chronological Age})$$

There has been a great deal of controversy in recent years over what intelligence tests measure. Many of the test items depend on either language or other specific cultural experiences for correct answers. Nevertheless, such tests can rather effectively predict school success. If school requires language and the tests measure language ability at a particular point of time in a child's life, then the test is a better-than-chance predictor of school performance.

Definition 10.58. $\mathcal{N}(0, 1)$ is the *standard* Gaussian (normal) distribution.

- We usually use Z to denote standard Gaussian RV.
- In Excel, use `NORMSINV(RAND())`.
In MATLAB, use `randn`.
- The standard normal cdf is denoted by $\Phi(z)$.
 - It inherits all properties of cdf.
 - Moreover, note that $\Phi(-z) = 1 - \Phi(z)$.

10.59. Relationship⁴⁷ between $\mathcal{N}(0, 1)$ and $\mathcal{N}(m, \sigma^2)$.

(a) An arbitrary Gaussian random variable with mean m and variance σ^2 can be represented as $\sigma Z + m$, where $Z \sim \mathcal{N}(0, 1)$.

This relationship can be used to generate general Gaussian RV from standard Gaussian RV.

(b) If $X \sim \mathcal{N}(m, \sigma^2)$, the random variable

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

is a standard normal random variable. That is, $Z \sim \mathcal{N}(0, 1)$.

- Creating a new random variable by this transformation is referred to as *standardizing*.
- The standardized variable is called “*standard score*” or “*z-score*”.

10.60. It is impossible to express the integral of a Gaussian PDF between non-infinite limits (e.g., (22)) as a function that appears on most scientific calculators.

⁴⁷To be proved in 10.81.

- An old but still popular technique to find integrals of the Gaussian PDF is to refer to tables that have been obtained by numerical integration.
 - An example of such table is Table 4, which lists $\Phi(z)$ for many values of positive z .
 - For $X \sim \mathcal{N}(m, \sigma^2)$, we can show that the CDF of X can be calculated from

$$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right).$$

Example 10.61. Suppose $Z \sim \mathcal{N}(0, 1)$. Evaluate the following probabilities.

(a) $P[-1 \leq Z \leq 1]$

(b) $P[-2 \leq Z \leq 2]$

Example 10.62. Suppose $X \sim \mathcal{N}(1, 2)$. Find $P[1 \leq X \leq 2]$.

Example 10.63. Signal Detection: Assume that in the detection of a digital signal, the background noise follows a normal distribution with a mean of 0 volt and standard deviation of 0.45 volt. The system assumes a digital 1 has been transmitted when the

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
0.00	0.5000	0.50	0.6915	1.00	0.8413	1.50	0.9332	2.00	0.97725	2.50	0.99379
0.01	0.5040	0.51	0.6950	1.01	0.8438	1.51	0.9345	2.01	0.97778	2.51	0.99396
0.02	0.5080	0.52	0.6985	1.02	0.8461	1.52	0.9357	2.02	0.97831	2.52	0.99413
0.03	0.5120	0.53	0.7019	1.03	0.8485	1.53	0.9370	2.03	0.97882	2.53	0.99430
0.04	0.5160	0.54	0.7054	1.04	0.8508	1.54	0.9382	2.04	0.97932	2.54	0.99446
0.05	0.5199	0.55	0.7088	1.05	0.8531	1.55	0.9394	2.05	0.97982	2.55	0.99461
0.06	0.5239	0.56	0.7123	1.06	0.8554	1.56	0.9406	2.06	0.98030	2.56	0.99477
0.07	0.5279	0.57	0.7157	1.07	0.8577	1.57	0.9418	2.07	0.98077	2.57	0.99492
0.08	0.5319	0.58	0.7190	1.08	0.8599	1.58	0.9429	2.08	0.98124	2.58	0.99506
0.09	0.5359	0.59	0.7224	1.09	0.8621	1.59	0.9441	2.09	0.98169	2.59	0.99520
0.10	0.5398	0.60	0.7257	1.10	0.8643	1.60	0.9452	2.10	0.98214	2.60	0.99534
0.11	0.5438	0.61	0.7291	1.11	0.8665	1.61	0.9463	2.11	0.98257	2.61	0.99547
0.12	0.5478	0.62	0.7324	1.12	0.8686	1.62	0.9474	2.12	0.98300	2.62	0.99560
0.13	0.5517	0.63	0.7357	1.13	0.8708	1.63	0.9484	2.13	0.98341	2.63	0.99573
0.14	0.5557	0.64	0.7389	1.14	0.8729	1.64	0.9495	2.14	0.98382	2.64	0.99585
0.15	0.5596	0.65	0.7422	1.15	0.8749	1.65	0.9505	2.15	0.98422	2.65	0.99598
0.16	0.5636	0.66	0.7454	1.16	0.8770	1.66	0.9515	2.16	0.98461	2.66	0.99609
0.17	0.5675	0.67	0.7486	1.17	0.8790	1.67	0.9525	2.17	0.98500	2.67	0.99621
0.18	0.5714	0.68	0.7517	1.18	0.8810	1.68	0.9535	2.18	0.98537	2.68	0.99632
0.19	0.5753	0.69	0.7549	1.19	0.8830	1.69	0.9545	2.19	0.98574	2.69	0.99643
0.20	0.5793	0.70	0.7580	1.20	0.8849	1.70	0.9554	2.20	0.98610	2.70	0.99653
0.21	0.5832	0.71	0.7611	1.21	0.8869	1.71	0.9564	2.21	0.98645	2.71	0.99664
0.22	0.5871	0.72	0.7642	1.22	0.8888	1.72	0.9573	2.22	0.98679	2.72	0.99674
0.23	0.5910	0.73	0.7673	1.23	0.8907	1.73	0.9582	2.23	0.98713	2.73	0.99683
0.24	0.5948	0.74	0.7704	1.24	0.8925	1.74	0.9591	2.24	0.98745	2.74	0.99693
0.25	0.5987	0.75	0.7734	1.25	0.8944	1.75	0.9599	2.25	0.98778	2.75	0.99702
0.26	0.6026	0.76	0.7764	1.26	0.8962	1.76	0.9608	2.26	0.98809	2.76	0.99711
0.27	0.6064	0.77	0.7794	1.27	0.8980	1.77	0.9616	2.27	0.98840	2.77	0.99720
0.28	0.6103	0.78	0.7823	1.28	0.8997	1.78	0.9625	2.28	0.98870	2.78	0.99728
0.29	0.6141	0.79	0.7852	1.29	0.9015	1.79	0.9633	2.29	0.98899	2.79	0.99736
0.30	0.6179	0.80	0.7881	1.30	0.9032	1.80	0.9641	2.30	0.98928	2.80	0.99744
0.31	0.6217	0.81	0.7910	1.31	0.9049	1.81	0.9649	2.31	0.98956	2.81	0.99752
0.32	0.6255	0.82	0.7939	1.32	0.9066	1.82	0.9656	2.32	0.98983	2.82	0.99760
0.33	0.6293	0.83	0.7967	1.33	0.9082	1.83	0.9664	2.33	0.99010	2.83	0.99767
0.34	0.6331	0.84	0.7995	1.34	0.9099	1.84	0.9671	2.34	0.99036	2.84	0.99774
0.35	0.6368	0.85	0.8023	1.35	0.9115	1.85	0.9678	2.35	0.99061	2.85	0.99781
0.36	0.6406	0.86	0.8051	1.36	0.9131	1.86	0.9686	2.36	0.99086	2.86	0.99788
0.37	0.6443	0.87	0.8078	1.37	0.9147	1.87	0.9693	2.37	0.99111	2.87	0.99795
0.38	0.6480	0.88	0.8106	1.38	0.9162	1.88	0.9699	2.38	0.99134	2.88	0.99801
0.39	0.6517	0.89	0.8133	1.39	0.9177	1.89	0.9706	2.39	0.99158	2.89	0.99807
0.40	0.6554	0.90	0.8159	1.40	0.9192	1.90	0.9713	2.40	0.99180	2.90	0.99813
0.41	0.6591	0.91	0.8186	1.41	0.9207	1.91	0.9719	2.41	0.99202	2.91	0.99819
0.42	0.6628	0.92	0.8212	1.42	0.9222	1.92	0.9726	2.42	0.99224	2.92	0.99825
0.43	0.6664	0.93	0.8238	1.43	0.9236	1.93	0.9732	2.43	0.99245	2.93	0.99831
0.44	0.6700	0.94	0.8264	1.44	0.9251	1.94	0.9738	2.44	0.99266	2.94	0.99836
0.45	0.6736	0.95	0.8289	1.45	0.9265	1.95	0.9744	2.45	0.99286	2.95	0.99841
0.46	0.6772	0.96	0.8315	1.46	0.9279	1.96	0.9750	2.46	0.99305	2.96	0.99846
0.47	0.6808	0.97	0.8340	1.47	0.9292	1.97	0.9756	2.47	0.99324	2.97	0.99851
0.48	0.6844	0.98	0.8365	1.48	0.9306	1.98	0.9761	2.48	0.99343	2.98	0.99856
0.49	0.6879	0.99	0.8389	1.49	0.9319	1.99	0.9767	2.49	0.99361	2.99	0.99861

Table 4: The standard normal CDF: $\Phi(z)$

voltage exceeds 0.9. (Otherwise, it assumes a digital 0 has been transmitted)

What is the probability of detecting a digital 1 when none was sent? [Montgomery and Runger, 2013, Ex. 4-15]

10.64. *Q-function*: $Q(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ corresponds to $P[Z > z]$ where $Z \sim \mathcal{N}(0, 1)$; that is $Q(z)$ is the probability of the “tail” of $\mathcal{N}(0, 1)$. The Q function is then a complementary cdf (ccdf).

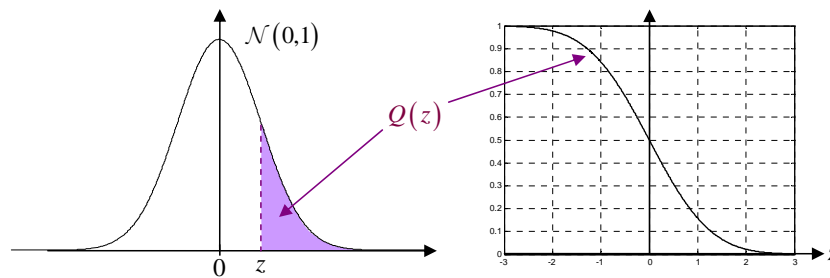


Figure 37: Q -function

- (a) Q is a decreasing function with $Q(0) = \frac{1}{2}$.
- (b) $Q(-z) = 1 - Q(z) = \Phi(z)$
- (c) Table 5 lists the values of $Q(z)$ for z between 3 to 5.
 - For z between 0 to 3, we use $Q(z) = 1 - \Phi(z)$.
 - For $z \geq 5$, the value of $Q(z)$ is extremely small. We may assume $Q(z) \approx 0$.

10.65. *Error function* (MATLAB): $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-x^2} dx = 1 - 2Q(\sqrt{2}z)$

- (a) It is an odd function of z .
- (b) For $z \geq 0$, it corresponds to $P[|X| < z]$ where $X \sim \mathcal{N}(0, \frac{1}{2})$.
- (c) $\lim_{z \rightarrow \infty} \text{erf}(z) = 1$

z	Q(z)	z	Q(z)	z	Q(z)	z	Q(z)	z	Q(z)
3.00	1.35E-03	3.40	3.37E-04	3.80	7.23E-05	4.20	1.33E-05	4.60	2.11E-06
3.01	1.31E-03	3.41	3.25E-04	3.81	6.95E-05	4.21	1.28E-05	4.61	2.01E-06
3.02	1.26E-03	3.42	3.13E-04	3.82	6.67E-05	4.22	1.22E-05	4.62	1.92E-06
3.03	1.22E-03	3.43	3.02E-04	3.83	6.41E-05	4.23	1.17E-05	4.63	1.83E-06
3.04	1.18E-03	3.44	2.91E-04	3.84	6.15E-05	4.24	1.12E-05	4.64	1.74E-06
3.05	1.14E-03	3.45	2.80E-04	3.85	5.91E-05	4.25	1.07E-05	4.65	1.66E-06
3.06	1.11E-03	3.46	2.70E-04	3.86	5.67E-05	4.26	1.02E-05	4.66	1.58E-06
3.07	1.07E-03	3.47	2.60E-04	3.87	5.44E-05	4.27	9.77E-06	4.67	1.51E-06
3.08	1.04E-03	3.48	2.51E-04	3.88	5.22E-05	4.28	9.34E-06	4.68	1.43E-06
3.09	1.00E-03	3.49	2.42E-04	3.89	5.01E-05	4.29	8.93E-06	4.69	1.37E-06
3.10	9.68E-04	3.50	2.33E-04	3.90	4.81E-05	4.30	8.54E-06	4.70	1.30E-06
3.11	9.35E-04	3.51	2.24E-04	3.91	4.61E-05	4.31	8.16E-06	4.71	1.24E-06
3.12	9.04E-04	3.52	2.16E-04	3.92	4.43E-05	4.32	7.80E-06	4.72	1.18E-06
3.13	8.74E-04	3.53	2.08E-04	3.93	4.25E-05	4.33	7.46E-06	4.73	1.12E-06
3.14	8.45E-04	3.54	2.00E-04	3.94	4.07E-05	4.34	7.12E-06	4.74	1.07E-06
3.15	8.16E-04	3.55	1.93E-04	3.95	3.91E-05	4.35	6.81E-06	4.75	1.02E-06
3.16	7.89E-04	3.56	1.85E-04	3.96	3.75E-05	4.36	6.50E-06	4.76	9.68E-07
3.17	7.62E-04	3.57	1.78E-04	3.97	3.59E-05	4.37	6.21E-06	4.77	9.21E-07
3.18	7.36E-04	3.58	1.72E-04	3.98	3.45E-05	4.38	5.93E-06	4.78	8.76E-07
3.19	7.11E-04	3.59	1.65E-04	3.99	3.30E-05	4.39	5.67E-06	4.79	8.34E-07
3.20	6.87E-04	3.60	1.59E-04	4.00	3.17E-05	4.40	5.41E-06	4.80	7.93E-07
3.21	6.64E-04	3.61	1.53E-04	4.01	3.04E-05	4.41	5.17E-06	4.81	7.55E-07
3.22	6.41E-04	3.62	1.47E-04	4.02	2.91E-05	4.42	4.94E-06	4.82	7.18E-07
3.23	6.19E-04	3.63	1.42E-04	4.03	2.79E-05	4.43	4.71E-06	4.83	6.83E-07
3.24	5.98E-04	3.64	1.36E-04	4.04	2.67E-05	4.44	4.50E-06	4.84	6.49E-07
3.25	5.77E-04	3.65	1.31E-04	4.05	2.56E-05	4.45	4.29E-06	4.85	6.17E-07
3.26	5.57E-04	3.66	1.26E-04	4.06	2.45E-05	4.46	4.10E-06	4.86	5.87E-07
3.27	5.38E-04	3.67	1.21E-04	4.07	2.35E-05	4.47	3.91E-06	4.87	5.58E-07
3.28	5.19E-04	3.68	1.17E-04	4.08	2.25E-05	4.48	3.73E-06	4.88	5.30E-07
3.29	5.01E-04	3.69	1.12E-04	4.09	2.16E-05	4.49	3.56E-06	4.89	5.04E-07
3.30	4.83E-04	3.70	1.08E-04	4.10	2.07E-05	4.50	3.40E-06	4.90	4.79E-07
3.31	4.66E-04	3.71	1.04E-04	4.11	1.98E-05	4.51	3.24E-06	4.91	4.55E-07
3.32	4.50E-04	3.72	9.96E-05	4.12	1.89E-05	4.52	3.09E-06	4.92	4.33E-07
3.33	4.34E-04	3.73	9.57E-05	4.13	1.81E-05	4.53	2.95E-06	4.93	4.11E-07
3.34	4.19E-04	3.74	9.20E-05	4.14	1.74E-05	4.54	2.81E-06	4.94	3.91E-07
3.35	4.04E-04	3.75	8.84E-05	4.15	1.66E-05	4.55	2.68E-06	4.95	3.71E-07
3.36	3.90E-04	3.76	8.50E-05	4.16	1.59E-05	4.56	2.56E-06	4.96	3.52E-07
3.37	3.76E-04	3.77	8.16E-05	4.17	1.52E-05	4.57	2.44E-06	4.97	3.35E-07
3.38	3.62E-04	3.78	7.84E-05	4.18	1.46E-05	4.58	2.32E-06	4.98	3.18E-07
3.39	3.49E-04	3.79	7.53E-05	4.19	1.39E-05	4.59	2.22E-06	4.99	3.02E-07

Table 5: The standard normal complementary CDF: $Q(z)$

(d) $\operatorname{erf}(-z) = -\operatorname{erf}(z)$

(e) $\Phi(x) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x}{\sqrt{2}} \right) \right) = \frac{1}{2} \operatorname{erfc} \left(-\frac{x}{\sqrt{2}} \right)$

(f) The complementary error function:

$$\operatorname{erfc}(z) = 1 - \operatorname{erf}(z) = 2Q(\sqrt{2}z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2} dx$$

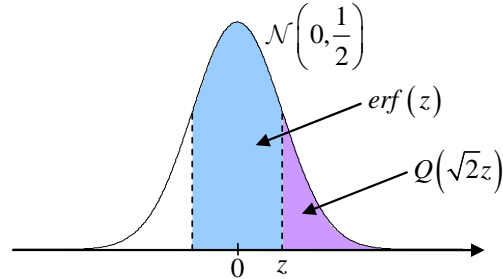


Figure 38: erf-function and Q -function

10.4.3 Exponential Distribution

Definition 10.66. The exponential distribution is denoted by $\mathcal{E}(\lambda)$.

(a) $\lambda > 0$ is a parameter of the distribution, often called the **rate parameter**.

(b) Characterized by

- $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0 \end{cases}$
- $F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0 \end{cases}$

(c) MATLAB:

- $X = \text{exprnd}(1/\lambda)$ or $\text{random}(\text{'exp'}, 1/\lambda)$
- $f_X(x) = \text{exp pdf}(x, 1/\lambda)$ or $\text{pdf}(\text{'exp'}, x, 1/\lambda)$
- $F_X(x) = \text{expcdf}(x, 1/\lambda)$ or $\text{cdf}(\text{'exp'}, x, 1/\lambda)$

10.67. The exponential distribution is intimately related to the Poisson process. In fact, the random variable X that equals the “distance” (or length or duration) between (any) successive events of a Poisson process with parameter λ is an exponential random variable with the same parameter.

Example 10.68. Exponential distribution is often used as a probability model for the (waiting) time until the next “rare” event occurs.

- time elapsed until the next earthquake in a certain region
- decay time of a radioactive particle
- time between independent events such as arrivals at a service facility or arrivals of customers in a shop.
- duration of a cell-phone call
- time it takes a computer network to transmit a message from one node to another.

10.69. In Example 10.37, we showed that $\mathbb{E}X = \frac{1}{\lambda}$.

Example 10.70. Suppose $X \sim \mathcal{E}(\lambda)$, find $P[1 < X < 2]$.

10.71. Survival-, survivor-, or reliability-function:

Exercise 10.72. Exponential random variable as a continuous version of geometric random variable: Suppose $X \sim \mathcal{E}(\lambda)$. Show that $\lfloor X \rfloor \sim \mathcal{G}_0(e^{-\lambda})$ and $\lceil X \rceil \sim \mathcal{G}_1(e^{-\lambda})$

Example 10.73. Phone Company A charges \$0.15 per minute for telephone calls. For any fraction of a minute at the end of a call, they charge for a full minute. Phone Company B also charges \$0.15 per minute. However, Phone Company B calculates its charge based on the exact duration of a call. If T , the duration of a call in minutes, is exponential with parameter $\lambda = 1/3$, what are the expected revenues per call $\mathbb{E}[R_A]$ and $\mathbb{E}[R_B]$ for companies A and B?

Solution: First, note that $\mathbb{E}T = \frac{1}{\lambda} = 3$. Hence,

$$\mathbb{E}[R_B] = \mathbb{E}[0.15 \times T] = 0.15\mathbb{E}T = \$0.45.$$

and

$$\mathbb{E}[R_A] = \mathbb{E}[0.15 \times \lceil T \rceil] = 0.15\mathbb{E}[\lceil T \rceil].$$

Now, recall, from Exercise 10.72, that $\lceil T \rceil \sim \mathcal{G}_1(e^{-\lambda})$. Hence, $\mathbb{E}[\lceil T \rceil] = \frac{1}{1-e^{-\lambda}} \approx 3.53$. Therefore,

$$\mathbb{E}[R_A] = 0.15\mathbb{E}[\lceil T \rceil] \approx 0.5292.$$

10.74. Memoryless property: The exponential r.v. is the only continuous⁴⁸ r.v. on $[0, \infty)$ that satisfies the memoryless property:

$$P[X > s + x | X > s] = P[X > x]$$

for all $x > 0$ and all $s > 0$ [18, p. 157–159]. In words, the future is independent of the past. The fact that it hasn't happened yet, tells us nothing about how much longer it will take before it does happen.

- Imagining that the exponentially distributed random variable X represents the lifetime of an item, the residual life of an item has the same exponential distribution as the original lifetime, regardless of how long the item has been already in use. In other words, there is no deterioration/degradation over time. If it is still currently working after 20 years of use, then today, its condition is “just like new”.

⁴⁸For discrete random variable, geometric random variables satisfy the memoryless property.

- In particular, suppose we define the set $B+x$ to be $\{x + b : b \in B\}$. For any $x > 0$ and set $B \subset [0, \infty)$, we have

$$P[X \in B + x | X > x] = P[X \in B]$$

because

$$\frac{P[X \in B + x]}{P[X > x]} = \frac{\int_{B+x} \lambda e^{-\lambda t} dt}{e^{-\lambda x}} \stackrel{\tau=t-x}{=} \frac{\int_B \lambda e^{-\lambda(\tau+x)} d\tau}{e^{-\lambda x}}.$$

Example 10.75. The exponential distribution is often used in reliability studies as the model for the time until failure of a device. For example, the lifetime of a semiconductor chip might be modeled as an exponential random variable with a mean of 40,000 hours.

The lack of memory property of the exponential distribution implies that the device does not wear out. That is, regardless of how long the device has been operating, the probability of a failure in the next 1000 hours is the same as the probability of a failure in the first 1000 hours of operation.

10.76. The lifetime L of a device with failures caused by random shocks might be appropriately modeled as an exponential random variable. However, the lifetime L of a device that suffers slow mechanical wear, such as bearing wear, is better modeled by other distributions such as the Weibull distribution.

10.77. Summary:

$X \sim$	Support S_X	$f_X(x) =$
Uniform $\mathcal{U}(a, b)$	(a, b)	$\begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{otherwise.} \end{cases}$
Normal (Gaussian) $\mathcal{N}(m, \sigma^2)$	\mathbb{R}	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$
Exponential $\mathcal{E}(\lambda)$	$(0, \infty)$	$\begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0 \end{cases}$

Table 6: Examples of probability density functions. Here, $\lambda, \sigma > 0$.