IES302 2011/2 Part I.7 Dr.Prapun

10 Continuous Random Variables

10.1 From Discrete to Continuous Random Variables

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 this type of random variable, the interesting fact is that

- any individual value has probability zero: $P[X = 1.7] = 0$

  $P[X = 5] = 0$

  $$P[X = x] = 0 \quad \text{for all } x \tag{22}$$

  and that

- the support is always uncountable

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

10.1. We can see from (22) 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 calculus[16] is required to formulate this continuous analog of a pmf.

[16] This is always a difficult concept for the beginning student.

Example 10.2. 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. [19, p 284]

Example 10.3. 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

This is just one example.

Example 10.4. The simplest example of a continuous random variable is the "random choice" of a number from the interval $(0, 1)$. $[0,1)$

The probability that the RV will occur in a region of size ℓ is the same for all region in $(0,1)$.

- In MATLAB, this can be generated by the command `rand`. In `Excel`, use `rand()`. Your calculator has a crude version of this : Ran#
- The generation is "unbiased" in the sense that "any number in the range is as likely to occur as another number."
- Again, because this is a continuous random variable, the probability that the randomly chosen number will take on a prespecified value is zero. [17]
- Histogram is flat over $[0, 1]$.

[17]So the above statement is true but not useful because it is true for all continuous random variables anyway. For any continuous random variable, the probability of a particular value x is 0 for any x; so they all have the same probability.

Definition 10.5. We say that X is a **continuous random variable**[18] if we can find a (real-valued) function f such that, for any set B, $P[X \in B]$ has the form

$$P[X \in B] = \int_B f_X(x)dx.$$

$$P[\text{some condition(s) on } X] = \int_{\text{all } x \text{ that satisfy the condition(s)}} f(x)\,dx$$

$$P[2 \leq X < 5] = P[2 < X < 5] = \int_2^5 f(x)dx \qquad P[X > 0] = \int_0^\infty f(x)dx$$

$$= P[2 \leq X \leq 5]$$

$$= P[2 < X \leq 5] \qquad P[X^2 > 1] = \int_{-\infty}^{-1} f(x)dx + \int_1^\infty f(x)dx$$

- In particular,

  $$P[a < X \leq b]$$
  $$= P[a < X < b]$$
  $$= P[a \leq X \leq b] = \int_a^b f_(x)dx.$$
  $$= P[a \leq X < b]$$

  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.

- Recall that when X is a discrete random variable,

  $$P[\text{some condition(s) on } X] = \sum_{\text{all } x \text{ that satisfy the condition(s)}} p_X(x)$$

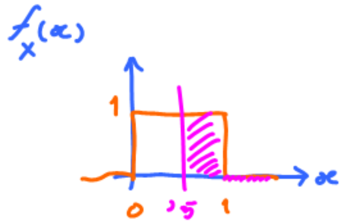
  $$P[2 < X < 5] = \sum_{2 < x < 5} p_X(x)$$

[18]To 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 (22) is satisfied. Alternatively, condition (22) 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.

$X = \text{rand}()$

Example 10.6. For the random variable generated by the `rand()` command in `Excel`,

$P[X > 0.5] = \int_{0.5}^{\infty} f_X(x)\,dx = \int_{0.5}^{1} 1\,dx = 0.5$



$P[X < 0.3] = 0.3$

Definition 10.7. 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 $\{x : f_X(x) > 0\}$.

10.2 Properties of PDF and CDF for Continuous Random Variables

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

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

$F_X(7)$

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

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

10.10. Given the cdf $F_X(x)$, we can find the pdf $f_X(x)$ by

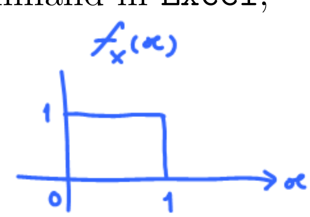
- If F_X is differentiable at x, we will set

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

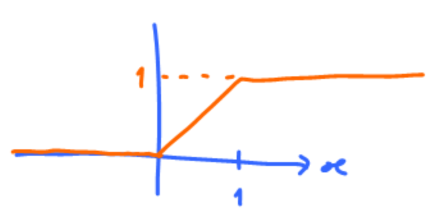
- If F_X is differentiable at x, we can set the values of $f_X(x)$ to be any value. Usually, the values are selected to give simple expression.

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

$F_X(-2)$



$\rightarrow$



Example 10.12. Suppose that the lifetime X of a device has the cdf

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

Observe that it is differentiable at each point x except for the two points $x = 0$ and $x = 2$. The probability density function is obtained by differentiation of the cdf which gives

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

In each of the finite number of points x at which F_X has no derivative, it does not matter what values we give f_X. These values do not affect $\int_B f_X(x)dx$. Usually, we give $f_X(x)$ the value 0 at any of these exceptional points.

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

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

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

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

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

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

10.16. f_X is nonnegative and $\int_{\mathbb{R}} f(x)dx = 1$.

Example 10.17. 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.

Theorem 10.18. Any nonnegative[19] function that integrates to one is a ***probability density function*** (pdf) [7, p. 139].

10.19. 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,

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

This is why we call f_X the density function.

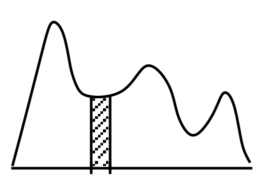


Figure 4: $P\left[x \le X \le x+\Delta x\right]$ is the area of the shaded vertical strip.

[19]or nonnegative a.e.

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

- In fact, $f_X(x) = \lim_{\Delta x \to 0} \frac{P[x < X \leq x + \Delta x]}{\Delta x}$
- 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.

10.20. Histogram and pdf [19, p 143 and 145]:

(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, you 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 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 histogram will begin to look more and more like the pdf.

(c) Conclusion: A probability density function can be seen as a "smoothed out" version of a probability histogram

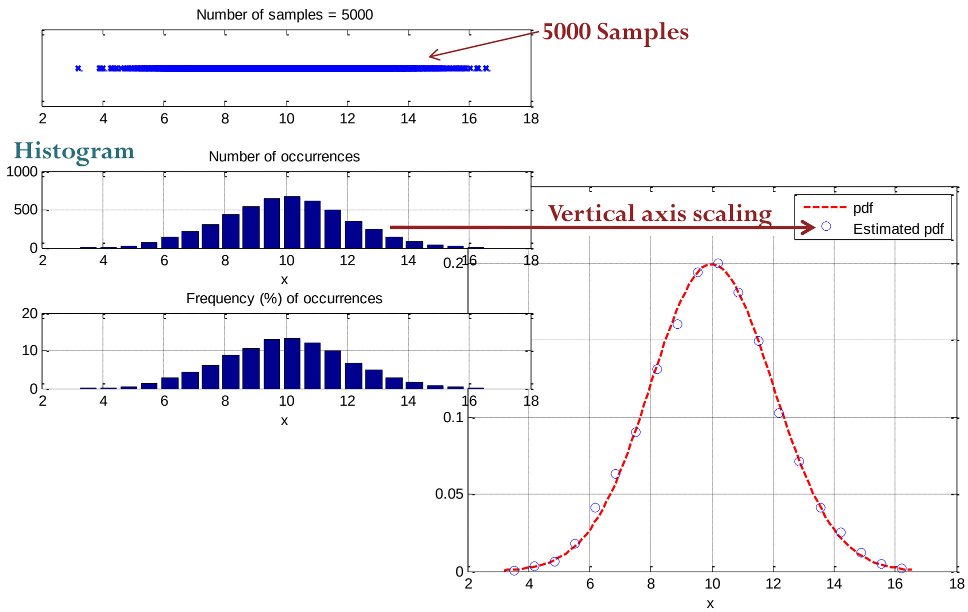


Figure 5: From histogram to pdf.

10.3 Expectation and Variance

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

Discrete RV

$\mathbb{E}X = \sum_x x\, p_X(x)$

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

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

$\mathbb{E}[X^2] = \sum_x x^2 p_X(x)$

In particular,

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

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

$f_X(x) =$ 

Example 10.22. For the random variable generated by the `rand()` command in `Excel`,

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^1 x \cdot 1\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

$f_X = 0$ outside the interval $(0,1)$

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \int_0^1 x^2 \cdot 1\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

$$\operatorname{Var} X = \mathbb{E}\left[X^2\right] - (\mathbb{E}X)^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

$$\sigma_X = \sqrt{\operatorname{Var} X} = 1/\sqrt{12}$$

10.23. 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.

10.24. Intuition/interpretation: As $n \to \infty$, the average of n independent samples of X will approach $\mathbb{E}X$.

- This observation is known as the "Law of Large Numbers".

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

(a) For $c \in \mathbb{R}$, $\mathbb{E}[c] = c$

(b) For $c \in \mathbb{R}$, $\mathbb{E}[X + c] = \mathbb{E}X + c$ and $\mathbb{E}[cX] = c\mathbb{E}X$

(c) For constants a, b, we have $\mathbb{E}[aX + b] = a\mathbb{E}X + b$.

(d) $\mathbb{E}[\cdot]$ is a **linear** operator: $\mathbb{E}[aX + bY] = a\mathbb{E}X + b\mathbb{E}Y$.

  (i) Homogeneous: $\mathbb{E}[cX] = c\mathbb{E}X$

  (ii) Additive: $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$

  (iii) Extension:

  - $\mathbb{E}\left[\sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n c_i \mathbb{E}X_i$.

(e) $\operatorname{Var} X = \mathbb{E}\left[X^2\right] - (\mathbb{E}X)^2$

(f) $\operatorname{Var} X \geq 0$.

(g) $\operatorname{Var} X \leq \mathbb{E}\left[X^2\right]$.

(h) $\operatorname{Var}[aX + b] = a^2 \operatorname{Var} X$.

$\sigma_{aX+b} = |a|\, \sigma_X$

10.4 Families of Continuous Random Variables

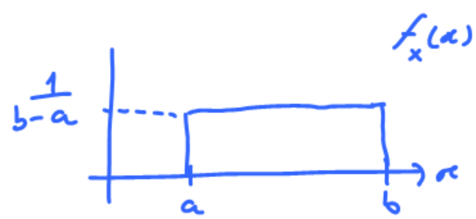
Theorem 10.18 states that any nonnegative function $f(x)$ whose integral over the interval $(-\infty, +\infty)$ equals 1 can be regarded as a probability density function of a random variable. In real-world applications, however, special mathematical forms naturally show up. In this section, we introduce 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 Uniform Distribution

Definition 10.26. For a uniform random variable on an interval $[a, b]$, we denote its family by uniform$([a, b])$ or $\mathcal{U}([a, b])$. Expressions that are synonymous with "X is a uniform random variable" are "X is uniformly distributed" and "X has a uniform distribution".

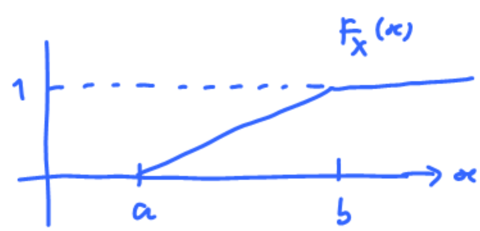
This family is characterized by

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



- The random variable X is just as likely to be near any value in $[a, b]$ as any other value.

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



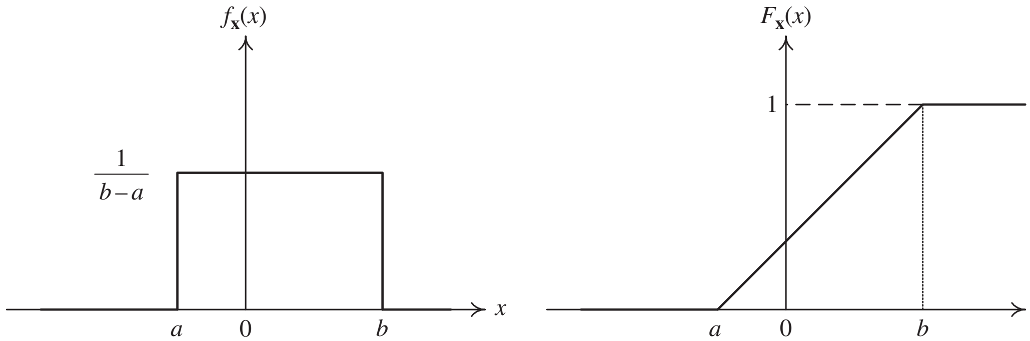
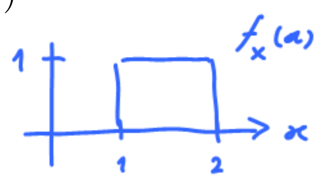


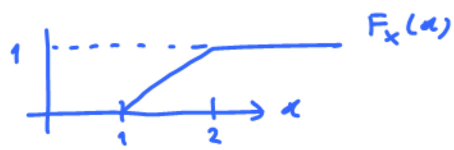
Figure 6: The pdf and cdf for the uniform random variable. [14, Fig. 3.5]

Example 10.27 (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.



10.28. 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.

10.4.2 Gaussian Distribution

10.29. 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 of them can be approximated by normal distribution.

Definition 10.30. ***Gaussian*** random variables:

(a) Often called ***normal*** random variables because they occur so frequently in practice

(Central Limit Theorem)

$\mathcal{N}(5, 4)$

mean (expected value) = 5

variance = 4 $\Rightarrow$ std. = 2

(b) Denoted by $\mathcal{N}\left(m, \sigma^2\right)$.

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

- In `Excel`, use `NORMSINV(RAND())`.
  In `MATLAB`, use `randn`.

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

- In `Excel`, use `NORMDIST(x,m,sigma,FALSE)`.
  In `MATLAB`, use `normpdf(x,m,sigma)`.
- Figure 9 displays the famous ***bell-shaped*** graph of the Gaussian pdf. This curve is also called the *normal* curve.
- Advanced calculus is required to prove that the area under the graph is indeed 1.

(e) $F_X(x) =$ `normcdf(x,m,sigma)`.

- In `Excel`, use `NORMDIST(x,m,sigma,TRUE)`.
  In `MATLAB`, use `normcdf(x,m,sigma)`.
- The standard normal cdf is sometimes denoted by $\Phi(x)$. It inherits all properties of cdf. Moreover, note that $\Phi(-x) = 1 - \Phi(x)$.
- If X is a $\mathcal{N}(m, \sigma^2)$ random variable, the CDF of X is

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

- It is impossible to express the integral of a Gaussian PDF between non-infinite limits as a function that appears on most scientific calculators.
  - 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.
    - One such table is the table that lists $\Phi(z)$ for many values of positive z.

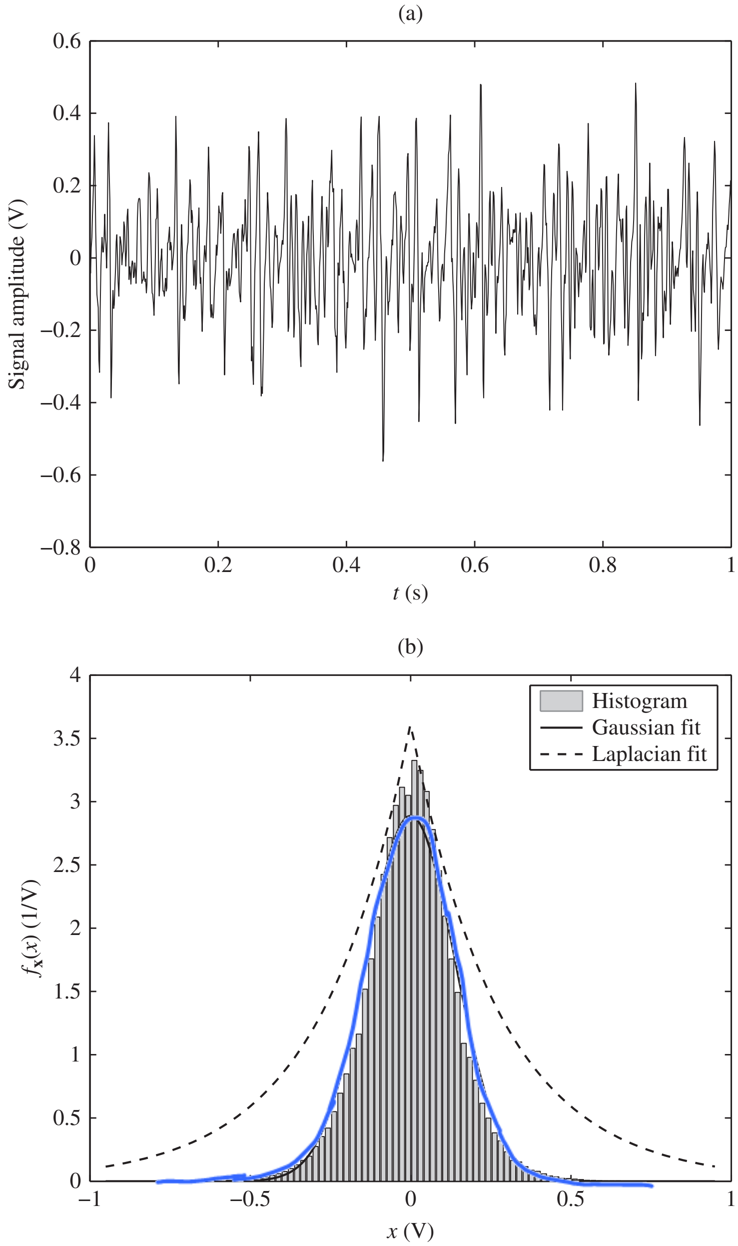


Figure 7: Electrical activity of a skeletal muscle: (a) A sample skeletal muscle (emg) signal, and (b) its histogram and pdf fits. [14, Fig. 3.14]

Question: Suppose $X \sim \mathcal{N}(1,2)$

$P[X > 3]$

$P[0 < X < 3]$

$m = 0$

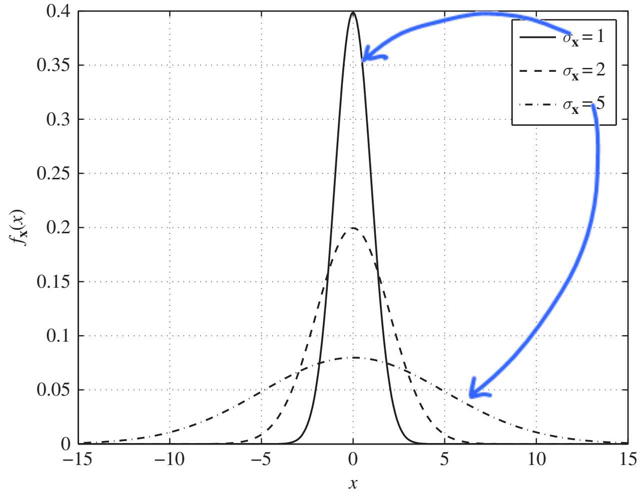


Figure 8: Plots of the zero-mean Gaussian pdf for different values of standard deviation, σ_X. [14, Fig. 3.15]

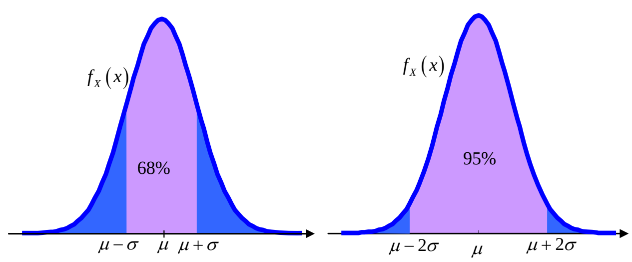


Figure 9: Probability density function of $X \sim \mathcal{N}(\mu, \sigma^2)$.

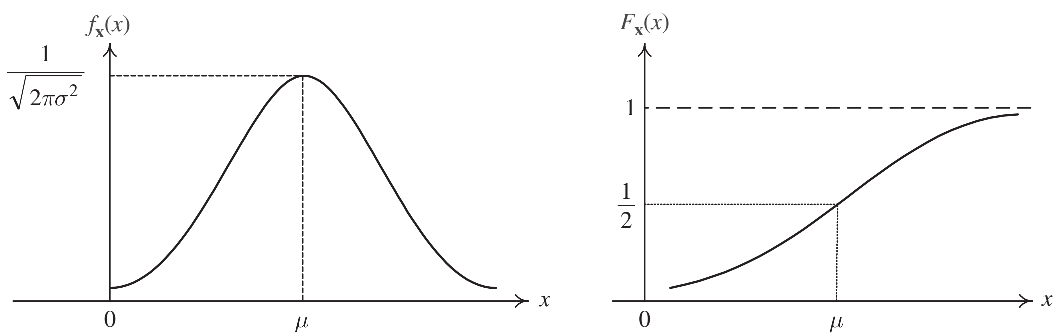


Figure 10: The pdf and cdf of $\mathcal{N}(\mu, \sigma^2)$. [14, Fig. 3.6]

10.31. $\mathbb{E}X = m$ and $\operatorname{Var} X = \sigma^2$.

10.32. $P[|X-\mu| < \sigma] = 0.6827$;
$P[|X-\mu| > \sigma] = 0.3173$;
$P[|X-\mu| > 2\sigma] = 0.0455$;
$P[|X-\mu| < 2\sigma] = 0.9545$

10.33. Relationship 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 $Z \sim \mathcal{N}(0,1)$.

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

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

$\mathbb{E} Z = 0$
$\operatorname{Var} Z = 1$

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.
- It is the key step to calculating a probability for an arbitrary normal random variable.

10.4.3 Exponential Distribution

Definition 10.34. 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

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