Jinguo Lian
Math437 Notes
January 15, 2016

3 Continuous Random Variables

Remember that discrete random variables can take only a countable number of possible values. On the other hand, a continuous random variable X has a range in the form of an interval or a union of non-overlapping intervals on the real line (possibly the whole real line). Also, for any $x \in \mathbb{R}, P(X = x) = 0$. Thus, we need to develop new tools to deal with continuous random variables. The good news is that the theory of continuous random variables is completely analogous to the theory of discrete random variables. Indeed, if we want to oversimplify things, we might say the following: take any formula about discrete random variables, and then replace sums with integrals, and replace PMFs with probability density functions (PDFs), and you will get the corresponding formula for continuous random variables. Of course, there is a little bit more to the story and that's why we need a chapter to discuss it.

3.1 Continuous Random Variables and their Distributions

3.1.1 The Cumulative Distribution Function (CDF)

Example 1. *I choose a real number uniformly at random in the interval $[a, b]$, and call it X. By uniformly at random, we mean all intervals in $[a, b]$ that have the same length must have the same probability. Find the CDF of X.*

Solution: *The uniformity implies that the probability of an interval of length l in [a,b] must be proportional to its length:*

$$P(X \in [x_1, x_2]) \propto (x_2 - x_1), \qquad \text{where } a \leq x_1 \leq x_2 \leq b.$$

Since $P(X \in [a, b]) = 1$, we conclude

$$P(X \in [x_1, x_2]) = \frac{x_2 - x_1}{b - a}, \qquad \text{where } a \leq x_1 \leq x_2 \leq b.$$

Now, let us find the CDF. By definition $F_X(x) = P(X \leq x)$, thus we immediately have

$$F_X(x) = 0, \text{ for } x < a$$

,

$$F_X(x) = 1, \ \textbf{for } x \geq b$$

. For $a \leq x \leq b$, we have

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(X \in [a, x]) \\ &= \frac{x - a}{b - a}. \end{aligned}$$

$$F_X(x) = \begin{cases} 0 & \textit{for } x < a \\ \frac{x-a}{b-a} & \textit{for } a \le x \le b \\ 1 & \textit{for } x > b \end{cases} \qquad (4.1) \qquad (1)$$

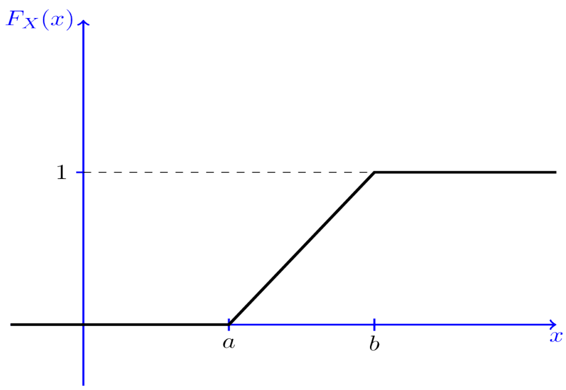

Figure 1: CDF for a continuous random variable uniformly distributed over [a,b].

One big difference that we notice here as opposed to discrete random variables is that the CDF is a continuous function, i.e., it does not have any jumps. Remember that jumps in the CDF correspond to points x for which $P(X = x) > 0$. Thus, the fact that the CDF does not have jumps is consistent with the fact that $P(X = x) = 0$ for all x. Indeed, we have the following definition for continuous random variables.

Definition 2. *A random variable X with CDF $F_X(x)$ is said to be continuous if $F_X(x)$ is a continuous function for all $x \in \mathbb{R}$.*

3.1.2 Probability Density Function (PDF)

To determine the distribution of a discrete random variable we can either provide its PMF or CDF. For continuous random variables, the CDF is well-defined so we can provide the CDF. However, the PMF does not work for continuous random variables, because for a continuous random variable $P(X = x) = 0$ for all $x \in R$. Instead, we define the probability density function (PDF). The PDF is the density of probability rather than the probability mass. The concept is very similar to mass density in physics: its unit is probability per unit length. To get a feeling for PDF, consider a continuous random variable X and define the function $f_X(x)$ as follows (wherever the limit exists):

$$f_X(x) = \lim_{\Delta \to 0^+} \frac{P(x < X \le x + \Delta)}{\Delta}.$$

The function $f_X(x)$ gives us the probability density at point x. It is the limit of the probability of the interval $(x, x + \delta]$ divided by the length of the interval as the length of the interval goes to 0. Remember that

$$P(x < X \le x + \Delta) = F_X(x + \Delta) - F_X(x).$$

so we conclude

$$\begin{aligned} f_X(x) &= \lim_{\Delta \to 0} \frac{F_X(x+\Delta) - F_X(x)}{\Delta} \\ &= \frac{dF_X(x)}{dx} = F'_X(x), \qquad \text{if } F_X(x) \text{ is differentiable at } x. \end{aligned}$$

Definition 3. *Consider a continuous random variable X with CDF $F_X(x)$. The function $f_X(x)$ defined by*

$$f_X(x) = \frac{dF_X(x)}{dx} = F'_X(x), \qquad \textit{if } F_X(x) \textit{ is differentiable at } x$$

is called the probability density function (PDF) of X.

Since the PDF is the derivative of the CDF, the CDF can be obtained from PDF by integration:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du.$$

Also, we have

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(u)du.$$

In particular, if we integrate over the entire real line, we must get 1, i.e.,

$$\int_{-\infty}^{\infty} f_X(u)du = 1.$$

Remark 4. *Consider a continuous random variable X with PDF $f_X(x)$. We have*

1. $f_X(x) \geq 0$ *for all* $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f_X(u)du = 1$.
3. $P(a < X \geq b) = F_X(b) - F_X(a) = \int_a^b f_X(u)du$.
4. *More generally, for a set A,* $P(X \in A) = \int_A f_X(u)du$.

Example 5. *An example of set A could be a union of some disjoint intervals. For example, if you want to find $P(\in [0,1] \cup [3,4])$, you can write*

$$P(X \in [0,1] \cup [3,4]) = \int_0^1 f_X(u)du + \int_3^4 f_X(u)du.$$

Example 6. *Let X be a continuous random variable with the following PDF*

$$f_X(x) = \begin{cases} ce^{-x} & x \geq 0 \\ 0 & \textit{otherwise} \end{cases}$$

where c is a positive constant.

1. *Find c.*
2. *Find the CDF of X, $F_X(x)$.*
3. *Find $P(1 < X < 3)$.*

3.1.3 The Mode, The Median, and Percentiles

Definition 7. *1. The mode of a continuous random variable X is the value of x for which the density function $f(x)$ is a maximum.*

2. The median m of a continuous random variable X is the solution of the equation

$$F(m) = P(X \leq m) = .50$$

3. Let X be a continuous random variable and $0 \leq p \leq 1$. The $100p^{th}$ of X is the number x_p defined by

$$F(x_p) = p$$

3.1.4 Expected Value and Variance

As we mentioned earlier, the theory of continuous random variables is very similar to the theory of discrete random variables. In particular, usually summations are replaced by integrals and PMFs are replaced by PDFs. The proofs and ideas are very analogous to the discrete case, so sometimes we state the results without mathematical derivations for the purpose of brevity.

Definition 8. *Remember that the expected value of a discrete random variable can be obtained as*

$$E(X) = \sum_{x_k \in R_X} x_k P_X(x_k).$$

Definition 9. *Now, by replacing the sum by an integral and PMF by PDF, we can write the definition of expected value of a continuous random variable as*

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

Example 10. *Let $X \sim Uniform(a, b)$. Find $E(X)$.*

Solution:

As we saw, the PDF of X is given by

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

so to find its expected value, we can write

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_a^b x(\frac{1}{b-a})dx \\ &= \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b dx = \frac{a+b}{2}. \end{aligned}$$

This result is intuitively reasonable: since X is uniformly distributed over the interval [a,b], we expect its mean to be the middle point, i.e., $E(X) = \frac{a+b}{2}$.

Example 11. *Let X be a continuous random variable with PDF*

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$$

Find the expected value of X.

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \frac{2}{3}.$$

Definition 12. *Remember the law of the unconscious statistician (LOTUS) for discrete random variables:*

$$E[g(X)] = \sum_{x_k \in R_X} g(x_k) P_X(x_k)$$

Law of the unconscious statistician (LOTUS) for continuous random variables:

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

As we have seen before, expectation is a linear operation, thus we always have

$$E[aX + b] = aE(X) + b$$

$$E[X_1 + X_2 + ... + X_n] = E(X_1) + E(X_2) + ... + E(X_n)$$

, for any set of random variables $X_1, X_2, ..., X_n$.

Definition 13.

$$Var(X) = E\big[(X - \mu_X)^2\big] = EX^2 - (EX)^2.$$

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

$$= EX^2 - (EX)^2 = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2$$

Also remember that for $a, b \in \mathbb{R}$, *we always have*

$$Var(aX + b) = a^2\, Var(X).$$

3.1.5 Functions of Continuous Random Variables

If X is a continuous random variable and $Y = g(X)$ is a function of X, then Y itself is a random variable. Thus, we should be able to find the CDF and PDF of Y. It is usually more straightforward to start from the CDF and then to find the PDF by taking the derivative of the CDF.

Remark 14. *The Method of Transformations:*

If we are interested in finding the PDF of $Y = g(X)$, *and the function g satisfies some properties, it might be easier to use a method called the method of transformations. Let's start with the case where g is a function satisfying the following properties:*

1. $g(x)$ is differentiable;

2. $g(x)$ is a strictly increasing function, that is, if $x_1 < x_2$, then $g(x_1) < g(x_2)$.

Note that since g is strictly increasing, its inverse function g^{-1} is well defined. That is, for each $y \in R_Y$, there exists a unique x_1 such that $g(x_1) = y$. We can write $x_1 = g^{-1}(y)$.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(g(X) \leq y) \\ &= P(X < g^{-1}(y)) \text{ since } g \text{ is strictly increasing} \\ &= F_X(g^{-1}(y)). \end{aligned}$$

To find the PDF of Y, we differentiate

$$f_Y(y) = \frac{d}{dy} F_X(x_1) \text{ where } g(x_1) = y$$

$$= \frac{dx_1}{dy} F'_X(x_1)$$

$$= f_X(x_1) \frac{dx_1}{dy}$$

$$= \frac{f_X(x_1)}{g'(x_1)} \text{ since } \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

We can repeat the same argument for the case where g is strictly decreasing. In that case, $g(x_1)$ will be negative, so we need to use $|g(x_1)|$. Thus, we can state the following theorem for a strictly monotonic function. (A function $g : R \to R$ is called strictly monotonic if it is strictly increasing or strictly decreasing.)

Theorem 15. *Suppose that X is a continuous random variable and $g : R \to R$ is a strictly monotonic differentiable function. Let $Y = g(X)$. Then the PDF of Y is given by*

$$f_Y(y) = \begin{cases} \frac{f_X(x_1)}{|g'(x_1)|} = f_X(x_1).|\frac{dx_1}{dy}| & \text{where } g(x_1) = y \\ 0 & \text{if } g(x) = y \text{ does not have a solution} \end{cases} \tag{4.5}$$

Theorem 16. *Consider a continuous random variable X with domain R_X, and let $Y = g(X)$. Suppose that we can partition R_X into a finite number of intervals such that $g(x)$ is strictly monotone and differentiable on each partition. Then the PDF of Y is given by*

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|} = \sum_{i=1}^{n} f_X(x_i). \left| \frac{dx_i}{dy} \right|$$

where $x_1, x_2, ..., x_n$ are real solutions to $g(x) = y$.

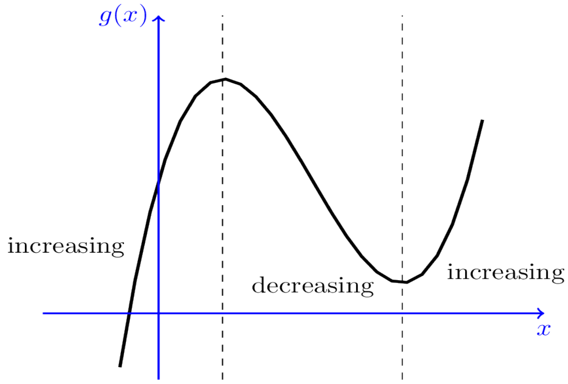

Figure 2: Partitioning a function to monotone parts.

Example 17. *Let X be a continuous random variable with PDF*

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \text{for all } x \in \mathbb{R}$$

and let $Y = X^2$. Find $f_Y(y)$.

Solution: *We note that the function $g(x) = x^2$ is strictly decreasing on the interval $(-\infty, 0)$, strictly increasing on the interval $(0, \infty)$, and differentiable on both intervals, $g(x) = 2x$. Thus, we can use the previous Equation. First, note that $R_Y = (0, \infty)$. Next, for any $y \in (0, \infty)$ we have two solutions for $y = g(x)$, in particular,*

$$x_1 = \sqrt{y}, \quad x_2 = -\sqrt{y}$$

Note that although $0 \in R_X$ it has not been included in our partition of R_X. This is not a problem, since $P(X = 0) = 0$. Indeed, in the statement of the previous Theorem, we could replace R_X by $R_X - A$, where A is any set for which $P(X \in A) = 0$. In particular, this is convenient when we exclude the endpoints of the intervals. Thus, we have

$$\begin{aligned}
f_Y(y) &= \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|} \\
&= \frac{f_X(\sqrt{y})}{|2\sqrt{y}|} + \frac{f_X(-\sqrt{y})}{|-2\sqrt{y}|} \\
&= \frac{1}{2\sqrt{2\pi y}} e^{-\frac{y}{2}} + \frac{1}{2\sqrt{2\pi y}} e^{-\frac{y}{2}} \\
&= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \text{ for } y \in (0, \infty).
\end{aligned}$$

3.2 Commonly Used Continuous Distributions

3.2.1 Uniform Distribution

Definition 18. *A continuous random variable X is said to have a Uniform distribution over the interval [a,b], shown as $X \sim Uniform(a,b)$, if its PDF is given by*

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

Remark 19.

$$E(X) = \frac{a+b}{2}$$

$$E(X^2) = \int_a^b x f_X(x) dx = \frac{a^2 + ab + b^2}{3}$$

3.2.2 Uniform Random Variable for Lifetimes; Survival Functions

Example 20. *Let T be the time from birth until death of a random selected member of a population. Assume that T has a uniform distribution on [0, 100]. Then*

$$f(t) = \begin{cases} \frac{1}{100}, & 0 \le t \le 100; \\ 0, & otherwise. \end{cases}$$

and

$$F(t) = \begin{cases} 0, & t¡0; \\ \frac{t}{100}, & 0 \le t \le 100; \\ 1, & t > 100. \end{cases}$$

The function $F(t)$ gives us the probability that the person dies by age t.

For example, the probability of death by age 57 is

$$P(T \le 57) = F(57) = \frac{57}{100} = .57$$

we might wish to find the probability that we survive beyond age 57. This is simply the probability that we do not die by age 57.

$$P(T > 57) = 1 - F(57) = 1 - .57 = .43$$

Definition 21. *The survival function is*

$$S(t) = P(T > t) = 1 - F(t)$$

3.2.3 A conditional Probability Involving the uniform Distribution

Example 22. *Let T be the lifetime random variable in , where T is uniform on [0,100]. Find*

1. $P(T \geq 50|T \geq 20)$

   Solution:

$$P(T \geq 50|T \geq 20) = \frac{P(T \geq 50 \ and \ P \geq 20)}{P(T \geq 20)}$$
$$= \frac{P(T \geq 50)}{P(T \geq 20)} = \frac{0.5}{0.8} = .625$$

2. $P(T \geq x|T \geq 20)$ *for* $x \in [20, 100]$.

   Solution:

$$P(T \geq x|T \geq 20) = \frac{P(T \geq x \ and \ P \geq 20)}{P(T \geq 20)}$$
$$= \frac{T \geq x}{P(T \geq 20)} = \frac{1 - \frac{x}{100}}{0.8} = \frac{100 - x}{80}$$

3.2.4 Exponential Distribution

Definition 23. *A continuous random variable* X *is said to have an exponential distribution with parameter* $\lambda > 0$*, shown as* $X \sim Exponential(\lambda)$*, if its PDF is given by*

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

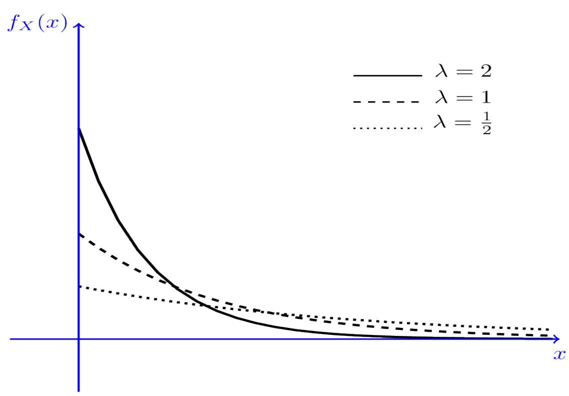

Figure 3: PDF of the exponential random variable.

Definition 24. *The cumulative distribution function of the exponential random variable:*

$$P(T \leq t) = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{\lambda t} \text{ for } t \geq 0$$

and the survival function

$$S(t) = 1 - F(t) = e^{\lambda t} \text{ for } t \geq 0$$

Remark 25. *If* $X \sim Exponential(\lambda)$*, then* $E(X) = \frac{1}{\lambda}$ *and* $Var(X) = \frac{1}{\lambda^2}$*.*

Example 26. *A company is studying the reliability of a part in a machine. The time T (in hours) from installation to failure of the part is a random variable. The study shows that T follows an exponential distribution with $\lambda = .001$. The probability that a part fails within 200 hours is*

$$P(0 \leq T \leq 200) = \int_0^{200} .001e^{-.001x}dx$$
$$= -e^{-.001x}|_0^{200} = 1 - e^{-.2} \approx 0.181$$

The probability that the part lasts for more than 500 hours

$$S(500) = e^{-0.5} \approx .607$$

Definition 27. *Let T be a random variable with density function $f(t)$ and cumulative distribution function $F(t)$. The **failure rate function** $\lambda(t)$ is defined by*

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}$$

The failure rate can be defined for any random variable, but is simplest to understand for an exponential variable. In this case, $\lambda(t) = \lambda$.

Remark 28. *Why the waiting time is exponential for events whose number follows a Poisson distribution?*

Assumption: *If the number of events in a time period of length 1 is poisson random variable with parameter λ, then the number of events in a time period of length t is a Poisson random variable with parameter λt.*

This is a reasonable assumption. For example, if the number of accident in a month at an intersection is Poisson variable with rate parameter $\lambda = 2$, then the assumption says that accident in a two month period will be Poisson with rate parameter of $2\lambda = 4$. Use this assumption, the probability of no accidents in an interval of length t is

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

However, there are no accidents in an interval of length t if and only if the waiting time T for the next accident is greater than t. Thus

$$P(X = 0) = P(T > t) = S(t) = e^{\lambda t}$$

3.2.5 A Conditional Probability Problem Involving the Exponential Distribution

Example 29. *Let T be the time to failure of the machine part. Where T is exponential with $\lambda = .001$. Find each of*

(a) $P(T \geq 150|T \geq 100)$ and (b) $P(T \geq x + 100|T \geq 100)$, for x in $[0, \infty)$.

Solution:

$$P(T \geq 150|T \geq 100) = \frac{P(T \geq 150 \text{ and } T \geq 100)}{P(T \geq 100)}$$

$$= \frac{P(T \geq 150)}{P(T \geq 100)} = \frac{e^{-.001(150)}}{e^{-.001(100)}} = e^{-0.5} \approx .951$$

if x is any real number in in the interval $[0, \infty)$,

$$P(T \geq x + 100 | T \geq 100) = \frac{P(T \geq x + 100 \text{ and } T \geq 100)}{P(T \geq 100)}$$

$$= \frac{P(T \geq x + 100)}{P(T \geq 100)} = \frac{e^{-.001(x+100)}}{e^{-.001(100)}} = e^{-0.01x}$$

From the point of view of waiting time until arrival of a customer, the memoryless property means that it does not matter how long you have waited so far. If you have not observed a customer until time a, the distribution of waiting time (from time a) until the next customer is the same as when you started at time zero. This is called memoryless.

3.2.6 Normal (Gaussian) Distribution

The normal distribution is the most widely-used of all the distributions. It can be used to model the distribution of heights, weights, test scores, measurement errors, stock portfolio returns, insurance portfolio loss etc.

Definition 30. *A continuous random variable Z is said to be a standard normal (standard Gaussian) random variable, shown as $Z \sim N(0, 1)$, if its PDF is given by*

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\}, \qquad \text{for all } z \in \mathbb{R}.$$

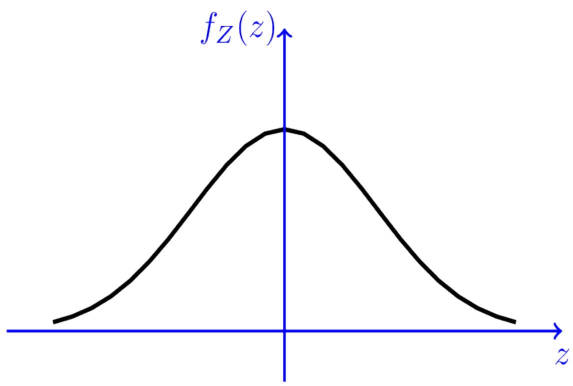

Figure 4: PDF of the standard normal random variable.

Remark 31. *If $Z \sim N(0, 1)$, then $E(Z) = 0$ and $Var(Z) = 1$.*

Remark 32. *The CDF of the standard normal distribution is denoted by the Φ function:*

$$\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{u^2}{2}\right\} du.$$

1. $\lim_{x\to\infty} \Phi(x) = 1, \quad \lim_{x\to-\infty} \Phi(x) = 0.$

2. $\Phi(0) = \frac{1}{2}.$

3. $\Phi(-x) = 1 - \Phi(x),$ *for all* $x \in \mathbb{R}$

Also, since the Φ function does not have a closed form, it is sometimes useful to use upper or lower bounds. In particular we can state the following bounds. For all $x \geq 0$,

$$\frac{1}{\sqrt{2\pi}} \frac{x}{x^2+1} \exp\left\{-\frac{x^2}{2}\right\} \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} \exp\left\{-\frac{x^2}{2}\right\}$$

Definition 33. *If Z is a standard normal random variable and $X = \sigma Z + \mu$, then X is a normal random variable with mean μ and variance σ^2, i.e,*

$$X \sim N(\mu, \sigma^2).$$

Remark 34. *If X is a normal random variable with mean μ and variance σ^2, i.e, $X \sim N(\mu, \sigma^2)$, then*

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

2. $F_X(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right),$

3. $P(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$

Theorem 35. *If $X \sim N(\mu_X, \sigma_X^2)$, and $Y = aX + b$, where $a, b \in \mathbb{R}$, then $Y \sim N(\mu_Y, \sigma_Y^2)$ where*

$$\mu_Y = a\mu_X + b, \quad \sigma_Y^2 = a^2\sigma_X^2.$$

Example 36. *Suppose we are looking at a national exam whose scores X are approximately normal with $\mu = 500$ and $\sigma = 100$. If we wish to find the probability that a score falls between 600 an 750, we must evaluate a difficult integral*

$$P(600 \leq X \leq 750) = \int_{600}^{750} \frac{1}{\sqrt{2\pi}100} e^{-\frac{x-500}{20000}} dx$$

Remark 37. *Normally, how to calculate the above integral, we use the following steps,*

1. *Linear transformation of normal random variables. For the above question, we can use $Z = \frac{x-\mu}{\sigma}$.*

2. *Transformation to a standard normal. $Z = \frac{x-\mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$.*

3. *Use Z-tables to find $\Phi(2.5) - \Phi(1) = 0.9938 - 0.8413 = .1525$*

Z-scores are useful for purpose other than table calculation. Z-value gives a distance from the mean in standard deviation units. Thus for the national exam with $\mu = 500$ and $\sigma = 100$, a student with an exam score of $x = 750$ and a transformed value of $z = 2.5$ can be described as being 2.5 standard deviation above the mean.

3.2.7 Sum of Independent, Identically Distributed Random Variables and Central Limit Theorem

Example 38. *Let the random variable X represent the loss on a single insurance policy. It was not normally distributed. Suppose we found that* $E(X) = \frac{1000}{3}$ *and* $Var(X) = \frac{500000}{9}$*, we also found that the probability for X. However, this information applies only to a single policy. If the company is willing to assume that all of the policies are independent and each is governed by the same distribution, the total claim loss S for the company is the sum of the losses on all the individual policies*

$$S = X_1 + X_2 + ... + X_{1000}$$

if we suppose that the company has 1000 policies.

Theorem 39. *Let* $X_1, X_2, ..., X_n$ *be independent random variables, all of which have the same probability distribution and thus the same mean* μ *and variance* σ^2*. If n is large, the sum*

$$S = X_1 + X_2 + ... + X_n$$

will be approximately normal with mean $n\mu$ *and variance* $n\sigma^2$*.*

the central limit theorem tell us the

$$\mu_S = 1000 \cdot \frac{1000}{3} = 333333.33, \sigma_S = 1000 \cdot \frac{500000}{9} = 7453.56$$

, then we use z-table to find the probability of total claim losses less than 350000 of the company,

$$P(S \leq 350000) = P(\frac{S - 333333.33}{7453.56} \leq \frac{350000 - 333333.33}{7453.56}) = P(Z \leq 2.24) = F_Z(2.24) = .9875$$

This result tell us the company is not likely to need more than 350000 to pay claims, which is helpful in planning. In general, the normal distribution is quite valuable because it applied in so many situation where independent and identical components are being added.

The central limit theorem enables us to understand why so may random variables are approximately normally distributed. This occures because many usefule random variables ate themselves sums of other independent random variables.

Remark 40. *If X is a normal random variable with mean* μ *and standard deviation* σ*, the we can easily find* x_p*, the* $100p^{th}$ *percentile of Z and the basic relationship of X and Z*

$$z_p = \frac{x_p - \mu}{\sigma} \to x_p = \mu + z_p\sigma$$

Example 41. *If X is a standard test score random variable with mean* $\mu = 500$ *and standard deviation* $\sigma = 100$*, then the* 99^{th} *percentile of X is*

$$x_{.99} = \mu + z_{.99}\sigma = 500 + 2.326(100) = 732.6$$

Remark 42. *Continuity correction is coved in basic statistics course. If you want to find the probability that a score was in the range from 600 to 700. without the continuity correction you would calculate*

$$P(500 \leq X \leq 700$$

With the continuity correction,

$$P(499.05 \leq X \leq 700.5)$$

3.2.8 Gamma Distribution

The gamma distribution is another widely used distribution. Its importance is largely due to its relation to exponential and normal distributions. The gamma random variable X can be used to model the waiting time for the n^{th} occurrences of events if successive occurrence are independent.

Definition 43. *The gamma function, shown by* $\Gamma(x)$*, is an extension of the factorial function to real (and complex) numbers. Specifically, if* $n \in \{1, 2, 3, ...\}$*, then*

$$\Gamma(n) = (n1)!$$

More generally, for any positive real number α*,* $\Gamma(\alpha)$ *is defined as*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x, \qquad \textit{for } \alpha > 0.$$

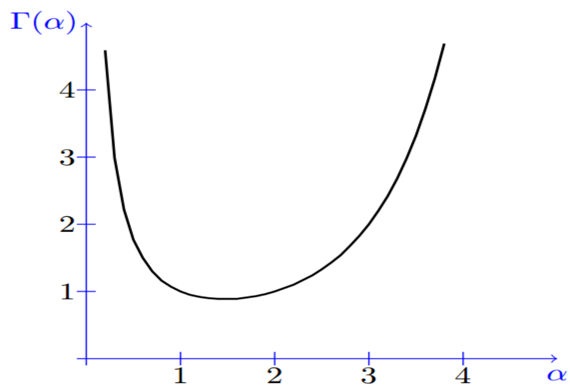

Figure 5: The Gamma function for some real values of α

Remark 44. *Properties of the gamma function*
For any positive real number α*:*

1. $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$
2. $\int_0^\infty x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}$*, for* $\lambda > 0$;
3. $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$;

4. $\Gamma(n) = (n-1)!$, *for* $n = 1, 2, 3, \cdots$;

5. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Definition 45. *A continuous random variable X is said to have a gamma distribution with parameters ¿0 and $\lambda > 0$, shown as $X \sim Gamma(\alpha, \lambda)$, if its PDF is given by*

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x > 0 \\ 0 & otherwise \end{cases}$$

If we let $\alpha = 1$, we conclude $Gamma(1, \lambda) = Exponential(\lambda)$. More generally, if you sum n independent $Exponential(\lambda)$ random variables, then you will get a $Gamma(n, \lambda)$ random variable.

Lemma 46. *Let $X_1, X_2, ..., X_n$ be independent random variables, all of which have the same exponential distribution with $f(x) = \lambda e^{-\lambda x}$. Then the sum $X_1 + X_2 + ... + X_n$ has a gamma distribution with parameters $\alpha = n$ and λ. Here X_1 is the waiting time for the first accident. X_2 is the waiting time between the first and second accidents. In general, X_i is the waiting time between accidents $i-1$ and i. Then*

$$\sum_{i=1}^{n} X_i = X$$

total waiting time for accident n.

Remark 47. *we find out that if $X \sim Gamma(\alpha, \lambda)$, then*

$$E(X) = \frac{\alpha}{\lambda}, \qquad Var(X) = \frac{\alpha}{\lambda^2}.$$

3.2.9 The Lognormal Distribution

Definition 48. *Random variable X has the lognormal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$, if $\ln(X)$ has the normal distribution with mean μ and standard deviation σ.*

Remark 49. *The parameter σ is the shape parameter of X while e^μ is the scale parameter of X.*

Equivalently, $X = e^Y$ where Y is normally distributed with mean μ and standard deviation σ. We can write $Y = \mu + \sigma Z$ where Z has the standard normal distribution. Hence we can write

$$X = e^{\mu + \sigma Z} = e^\mu (e^Z)^\sigma$$

Random variable e^Z has the lognormal distribution with parameters 0 and 1, and naturally enough, this is the standard lognormal distribution. The lognormal distribution is used to model continuous random quantities when the distribution is believed to be skewed, such as insurance claim severity or investment returns, certain income and lifetime variables.

Definition 50. *The probability density function of the lognormal distribution with parameters μ and σ is given by*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{[\ln(x)-\mu]^2}{2\sigma^2}\right), \quad x \in (0,\infty)$$

This function is difficult to work with, but we will not need it. We will use standard normal probability to find it. The lognormal distribution function F is given by

$$F(x) = \Phi\left[\frac{\ln(x)-\mu}{\sigma}\right], \quad x \in (0,\infty)$$

In particular, the mean and variance of X are

$$E(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$Var(X) = \exp\left[2(\mu+\sigma^2)\right] - \exp\left(2\mu+\sigma^2\right)$$

Example 51. *Supose the random variable Y is used as a model for claim amounts. We wish to find the probability of the occurrence of a claim greater than 1300. Since X is normal with $\mu = 7$ and $\sigma = 0.5$, we can use Z-tables. The probability of a claim less than or equal to 1300 is*

$$P(Y \le 1300) = P(e^X \le 1300) = P(X \le \ln 1300) = P(Z \le \frac{\ln 1300 - 7}{0.50}) = F_Z(.34) = .6331$$

Remark 52. *The Lognormal distribution for a stock price: the continuous growth model, value of asset at time t if growth is continuous at rate r*

$$A(t) = a(0)e^{rt}$$

3.2.10 The Pareto Distribution

The Pareto distribution can be used to model certain insurance loss amounts. The Pareto distribution has a number of different formulations. The one we have chosen involves two constants, α and β.

Definition 53. *1. Pareto density function $f(x) = \frac{\alpha}{\beta}(\frac{\beta}{x})^{\alpha+1}$ for $\alpha > 2, x \ge \beta > 0$.*

2. Pareto Cumulateive Distribution Function, $F(x) = 1 - (\frac{\beta}{x})^{\alpha}$ for $\alpha > 2, x \ge \beta > 0$.

3. $E(X) = \frac{\alpha\beta}{\alpha-1}$

4. $Var(X) = \frac{\alpha\beta^2}{\alpha-2} - (\frac{\alpha\beta}{\alpha-1})^2$

5. The failure rate of a Pareto random variable $\lambda(x) = \frac{\frac{\alpha}{\beta}(\frac{\beta}{x})^{\alpha+1}}{(\frac{\beta}{x})^{\alpha}} = \frac{\alpha}{x}$

3.2.11 The Weibull Distribution

Researchers who study units that fails or die often like to think in terms of the failure rate. They might decide to use an exponential distribution model if they believe the failure rate is constant. If they believe that the failure rate increases with time or age, then the Weibull distribution can provide a useful model.

Definition 54. 1. *Weibull Density function* $f(x) = \alpha\beta x^{\alpha-1}e^{-\beta x^{\alpha}}$ *for* $x \geq 0$, $\alpha > 0, \beta > 0$.

2. *Weibull Cumulative distribution Function* $F(x) = 1 - e^{-\beta x^{\alpha}}$, *for* $x \geq 0$, $\alpha > 0, \beta > 0$.

3. $E(X) = \frac{\Gamma(1+\frac{1}{\alpha})}{\beta^{\frac{1}{\alpha}}}$.

4. $Var(X) = \frac{1}{\beta^{\frac{2}{\alpha}}}\left[\Gamma(1+\frac{2}{\alpha}) - \Gamma(1+\frac{1}{\alpha}^2)\right]$.

5. *The failure rate:* $\lambda(x) = \alpha\beta(x^{\alpha-1})$.

3.2.12 The Beta distribution

The beta distribution is defined on the interval [0,1]. Thus the beta distribution can be used to model random variable whose outcomes are percent ranging from 0% to 100% and written in decimal form. It can be applied to study the percent of defective unites in a manufacturing process, the percent of errors made in data entry, the percent of clients satisfied with their service, and similar variables.

Definition 55. 1. *Beta density function:* $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ *for* $0 < x < 1$ *and* $\alpha > 0, \beta > 0$.

2. *When* $\alpha - 1$ *and* $\beta - 1$ *ate non-negative integers, the cummulative distribution function can be found by integrating a polynomial.*

3. *a useful identity* $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

4. $E(X) = \frac{\alpha}{\alpha+\beta}$.

5. $Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Example 56. *A management firm handles investment accounts for a large number of clients. The percent of clients who telephone the firm for information or service in a given month is beta random variable with* $\alpha = 4$ *and* $\beta = 3$*. Find the density function* $f(x) = 60(x^3 - 2x^4 + x^5)$ *for* $0 < x < 1$.

3.2.13 The Chi-Square Distribution

Definition 57. *1. The Chi-Square Density function:*

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, \quad x \in (0, \infty)$$

Where n is usually a positive integer, is the degree of freedom.

2. The distribution function:

$$F(x) = \frac{\Gamma(n/2, x/2)}{\Gamma(n/2)}, \quad x \in (0, \infty)$$

3. If X has the chi-square distribution with n degrees of freedom then $E(X) = n, Var(X) = 2n$.

4. X_n has the chi-square distribution with n degrees of freedom, then the distribution of the standard score

$$Z_n = \frac{X_n - n}{\sqrt{2n}}$$

3.2.14 Mixed Random Variables

Here, we will discuss mixed random variables. These are random variables that are neither discrete nor continuous, but are a mixture of both. In particular, a mixed random variable has a continuous part and a discrete part. Thus, we can use our tools from previous chapters to analyze them.

Remark 58. *In general, the CDF of a mixed random variable Y can be written as the sum of a continuous function and a staircase function:*

$$F_Y(y) = C(y) + D(y).$$

We differentiate the continuous part of the CDF. In particular, let's define

$$c(y) = \frac{dC(y)}{dy}, \text{ wherever } C(y) \text{ is differentiable.}$$

Note that this is not a valid PDF as it does not integrate to one. Also, let $\{y_1, y_2, y_3, ...\}$ be the set of jump points of $D(y)$, i.e., the points for which $P(Y = y_k) > 0$. We then have

$$\int_{-\infty}^{\infty} c(y)dy + \sum_{y_k} P(Y = y_k) = 1.$$

The expected value of Y can be obtained as

$$EY = \int_{-\infty}^{\infty} yc(y)dy + \sum_{y_k} y_k P(Y = y_k).$$

Example 59. *Let X be a continuous random variable with the following PDF:*

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & otherwise \end{cases}$$

Let also

$$Y = g(X) = \begin{cases} X & 0 \leq X \leq \frac{1}{2} \\ \frac{1}{2} & X > \frac{1}{2} \end{cases}$$

Find the CDF of Y.

Solution: *First we note that $R_X = [0,1]$. For $x \in [0,1]$, $0 \geq g(x) \geq \frac{1}{2}$. Thus, $R_Y = [0, \frac{1}{2}]$, and therefore*

$$F_Y(y) = 0, \qquad \text{for } y < 0,$$

$$F_Y(y) = 1, \qquad \text{for } y > \frac{1}{2}.$$

Now note that

$$P\left(Y = \frac{1}{2}\right) = P\left(X > \frac{1}{2}\right)$$

$$= \int_{\frac{1}{2}}^{1} 2x dx = \frac{3}{4}.$$

Also, for $0 < y < \frac{1}{2}$,

$$FY_(y) = P(Y \geq y)$$

$$= P(X \geq y)$$

$$= \int_0^y 2x dx$$

$$= y^2.$$

Thus, the CDF of Y is given by

$$F_Y(y) = \begin{cases} 1 & y \geq \frac{1}{2} \\ y^2 & 0 \leq y < \frac{1}{2} \\ 0 & otherwise \end{cases}$$

We note that the CDF is not continuous, so Y is not a continuous random variable. On the other hand, the CDF is not in the staircase form, so it is not a discrete random variable either. It is indeed a mixed random variable. There is a jump at $y = \frac{1}{2}$, and the amount of jump is $1\frac{1}{4} = \frac{3}{4}$, which is the probability that $Y = \frac{1}{2}$. The CDF is continuous at other points.

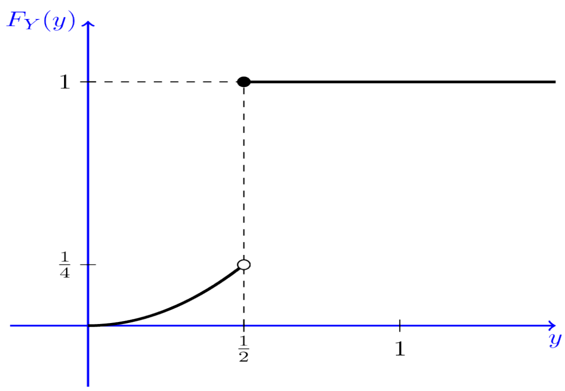

Figure 6: PMF of a Geometric(0.3) random variable.