Probability

Jan-April 2014

Chennai Mathematical Institute

Professor Nandini Kannan

Chapter 3

Continuous Random Variables

Definition: A random variable X is **absolutely continuous** (or just continuous) if its cdf is a continuous function of x.
We may also use the following definition:

A random variable is absolutely continuous if there exists a function $f, f \geq 0$ such that

$$F_X(x) = \int_{-\infty}^{x} f(t)dt \qquad \forall x \in \mathcal{R}. \tag{3.1}$$

$f(.)$ is called the probability density function (pdf).
Using the Fundamental Theorem of Calculus, if $f(.)$ is continuous, we have

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

We write $X \sim f_X(x)$.

Theorem 3.0.1. *A function $f_X(x)$ is a pdf iff*

(a) $f_X(x) \geq 0 \qquad \forall x;$

(b) $\int_{-\infty}^{\infty} f_X(x)dx = 1.$

Proof: The Necessity follows directly from the properties of F_X. The sufficiency requires us to show that if a function satisfies the two conditions above, then there exists a unique random variable for which this is the pdf.

■

If a random variable is continuous, then $P(X = x) = 0$ for all $x \in \mathcal{R}$.

3.1 Transformations

Let X be a random variable with cdf $F_X(x)$. Sometimes, we are interested not in X, but a transformed version of X. Let $Y = g(X)$ be a function of X.
Clearly, Y is also a random variable. We can describe the probabilistic behaviour of Y in terms of that of X. We have

$$P[Y \in A] = P[g(X) \in A] \text{ for any } A.$$

Let $\mathcal{X}$ be the sample space of X. Then

$$g : \mathcal{X} \to \mathcal{Y}$$

where $\mathcal{Y}$ is the set of possible values of the random variable Y.
We can define the inverse mapping

$$g^{-1} : \mathcal{Y} \to \mathcal{X}$$

where

$$g^{-1}(A) = \{x \in \mathcal{X}; g(x) \in A\}.$$

Therefore

$$\begin{aligned} P[Y \in A] &= P[g(X) \in A] \\ &= P[\{x \in \mathcal{X}; g(x) \in A\}] \\ &= P[X \in g^{-1}(A)]. \end{aligned}$$

This defines the probability distribution of Y. This distribution satisfies Kolmogorov's axioms.

If X and Y are continuous r.v.'s, it is possible to find simple formulae for the cdf and the the pdf of Y in terms of the cdf and pdf of X.
We have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P[\{x \in \mathcal{X}; g(x) \leq y\}] \\ &= \int_{\{x \in \mathcal{X}; g(x) \leq y\}} f_X(x)dx. \end{aligned}$$

This is called the **method of distribution functions**.

Example: Let X be a random variable with pdf

$$f_X(x) = \begin{cases} \frac{2x}{\pi^2}, & 0 < x < \pi; \\ 0, & \text{o.w.} \end{cases}$$

Let $Y = \sin X$. The range of Y is $(0, 1)$.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P[\sin X \leq Y] \\ &= P[0 \leq X \leq x_1] + P[x_2 \leq X \leq \pi]. \end{aligned}$$

where $x_1 = \sin^{-1} y, x_2 = \pi - \sin^{-1} y$.
Therefore

$$\begin{aligned} F_Y(y) &= \int_0^{x_1} \frac{2x}{\pi^2}dx + \int_{x_2}^{\pi} \frac{2x}{\pi^2}dx \\ &= \frac{x_1^2}{\pi^2} + 1 - \frac{x_2^2}{\pi^2}. \end{aligned}$$

■

Definition: The **support** of a random variable or distribution is the set of values for which the pdf (pmf) is nonzero:

$$\mathcal{X} = \{x : f_X(x) > 0\}.$$

Let $\mathcal{Y}$ be the sample space for Y. We have

$$\mathcal{Y} = \{y : y = g(x) \text{ for some } \ x \in \mathcal{X}\}.$$

If the transformation $x \to g(x)$ is monotone, we can obtain simple expressions for $F_Y(y)$.

Theorem 3.1.1. *Let $X \sim F_X(x)$ and $Y = g(X)$.*

(a) If g is an increasing function, then

$$F_Y(y) = F_X[g^{-1}(y)] \textit{ for } \ y \in \mathcal{Y}.$$

(b) If g is a decreasing function, then

$$F_Y(y) = 1 - F_X[g^{-1}(y)] \textit{ for } \ y \in \mathcal{Y}.$$

Proof: (a) If g is an increasing function, it is one-to-one and onto from $\mathcal{X} \to \mathcal{Y}$. In other words, each x goes to only one y, and each y comes from at most one x [one-to one], and for each $y \in \mathcal{Y}$, there is an $x \in \mathcal{X}$ such that $g(x) = y$ (onto). We have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P[\{x \in \mathcal{X}; g(x) \leq y\}] \\ &= P[\{x \in \mathcal{X}; x \leq g^{-1}(y)\}] \\ &= F_X[g^{-1}(y)]. \end{aligned}$$

(a) If g is a decreasing function, we have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \geq g^{-1}(Y)] \\ &= 1 - F_X[g^{-1}(y)]. \end{aligned}$$

■

Example: Let X be a random variable with pdf

$$f_X(x) = \begin{cases} e^{-x}, & x > 0; \\ 0, & \text{o.w.} \end{cases}$$

Let $Y = \ln X$. Therefore $g^{-1}(y) = e^y$. We have

$$\frac{d}{dx} g(x) = \frac{d}{dx} \ln x = \frac{1}{x} > 0.$$

Therefore g is an increasing function. As x ranges from 0 to ∞, y ranges from $-\infty$ to ∞. We have

$$\begin{aligned} F_Y(y) &= F_X[g^{-1}(y)] \\ &= F_X(e^y). \end{aligned}$$

We have

$$F_X(x) = \int_0^x e^{-t} dt = 1 - e^{-x}.$$

Substituting for $F_X(x)$, we have

$$F_Y(y) = 1 - e^{-e^y}.$$

■

Example: Let X be a random variable with pdf

$$f_X(x) = \begin{cases} \alpha x^{\alpha-1}, & 0 < x < 1; \\ 0, & \text{o.w.} \end{cases}$$

Let $Y = -\ln X$. Therefore $g^{-1}(y) = e^{-y}$. We have

$$\frac{d}{dx} g(x) = \frac{d}{dx}(-\ln x) = -\frac{1}{x} < 0.$$

Therefore g is a decreasing function. As x ranges from 0 to 1, y ranges from 0 to ∞. We have

$$F_X(x) = \int_0^x \alpha t^{\alpha-1} dt = x^\alpha.$$

$$\begin{aligned} F_Y(y) &= 1 - F_X[g^{-1}(y)] \\ &= 1 - F_X(e^{-y}) = 1 - e^{-\alpha y}. \end{aligned}$$

■

If the pdf of Y is continuous, it can be obtained by differentiating the cdf.

Theorem 3.1.2. *Let $X \sim f_X(x)$ and $Y = g(X)$, where g is a monotone function. Suppose $f_X(x)$ is continuous on $\mathcal{X}$ and $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$. Then*

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)] \left| \frac{d}{dy} g^{-1}(y) \right|, & y \in \mathcal{Y}; \\ 0, & o.w. \end{cases}$$

Proof: Use the chain rule.
■

Example Let X be a non-negative continuous random variable with pdf $f_X(x)$. Let $Y = X^\alpha, \alpha > 0$. Then

$$\frac{d}{dx} g(x) = \frac{d}{dx} x^\alpha = \alpha x^{\alpha-1} > 0,$$

which implies $g(.)$ is an increasing function. We have $g^{-1}(y) = y^{\frac{1}{\alpha}}$. Using the theorem, we have

$$f_Y(y) = \begin{cases} f_X\left[y^{\frac{1}{\alpha}}\right] \frac{1}{\alpha} y^{\frac{1}{\alpha}-1}, & y > 0; \\ 0, & \text{o.w.} \end{cases}$$

■

3.2 Expectation

Definition: Let X be a continuous random variable. The expected value or mean of the random variable $g(X)$ is

$$E[g(X)] = \int g(x) f_X(x) dx, \tag{3.2}$$

provided the integral exist. If $E|g(X)| = \infty$, we say that $E(g(X))$ **does not exist**.

Example: Let X be a continuous random variable with

$$f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2} \qquad x \in \mathcal{R}.$$

This is the Cauchy distribution. We have

$$E|X| = \int_{-\infty}^{\infty} \frac{|x|}{\pi} \frac{1}{1+x^2} dx = \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2} dx.$$

For any $M > 0$,

$$\int_0^M \frac{x}{1+x^2} dx = \frac{1}{2} \log(1+x^2) \Big|_0^M = \frac{1}{2} \log(1+M^2).$$

We have

$$E|X| = \lim_{M \to \infty} \frac{2}{\pi} \int_0^M \frac{x}{1+x^2} dx = \infty,$$

which implies $E(X)$ **does not exist**. ■

Note: The integral $\int_{-\infty}^{\infty} g(x) dx$ exists provided

$$\lim_{\substack{a \to \infty \\ b \to \infty}} \int_{-b}^{a} g(x) dx$$

exists. It is possible for $\lim_{a \to \infty} \int_{-a}^{a} g(x) dx$ to exist without the existence of $\int_{-\infty}^{\infty} g(x) dx$. For the previous example, we have

$$\lim_{a \to \infty} \int_{-a}^{a} \frac{x}{\pi} \frac{1}{1+x^2} = 0.$$

3.2.1 Properties of Expectation

Theorem 3.2.1. *Let X be a random variable and let a, b, c be constants. Then for any functions $g_1(X)$ and $g_2(X)$ whose expectations exist:*

1. $E[ag_1(X) + bg_2(X) + c] = aE[g_1(X)] + bE[g_2(X)] + c$.
2. *If* $g_1(x) \geq 0 \ \ \forall \ \ x$, *then*
$$E[g_1(X)] \geq 0.$$
3. *If* $g_1(x) \geq g_2(x) \ \ \forall \ \ x$, *then*
$$E[g_1(X)] \geq E[g_2(X)].$$

4. If $a \leq g_1(x) \leq b \;\; \forall \;\; x$, then

$$a \leq E[g_1(X)] \leq b.$$

Proof: Property 1. We have

$$\begin{aligned}
E[ag_1(X) + bg_2(X) + c] &= \int [ag_1(X) + bg_2(X) + c]f(x)dx \\
&= \int ag_1(X)f(x)dx + \int bg_2(X)f(x)dx + \int cf(x)dx \\
&= a\int g_1(X)f(x)dx + b\int g_2(X)f(x)dx + c\int f(x)dx \\
&= aE[g_1(X)] + bE[g_2(X)] + c.
\end{aligned}$$

This is the linearity property of the expectation operator.
Property 2. We have

$$E[g_1(X)] = \int g_1(x)f(x)dx \geq 0.$$

Property 3. We have

$$E[g_1(X)] = \int g_1(x)f(x)dx \geq \int g_2(x)f(x)dx = E[g_2(X)].$$

Property 4. We have

$$\begin{aligned}
&\int af(x)dx \leq \int g_1(x)f(x)dx \leq \int bf(x)dx \\
&\Rightarrow a \leq E[g_1(X)] \leq b.
\end{aligned}$$

If the random variables were discrete, the integrals would simply be replaced by summations.
■

3.2.2 Uniform or Rectangular Distribution

Definition: A random variable X is said to have a uniform distribution on the interval (a, b) if its pdf is given by

$$f(x|a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b; \\ 0, & \text{otherwise.} \end{cases}$$

We write $X \sim U(a, b)$

Theorem 3.2.2. *Let $X \sim U(a, b)$. Then*

$$\mu = \frac{a+b}{2} \tag{3.3}$$

and

$$\sigma^2 = \frac{(b-a)^2}{12}. \tag{3.4}$$

■

3.2.3 Normal or Gaussian Distribution

The most well known and widely used continuous random variable is the **Normal** random variable. The form of the distribution was discovered early in the history of probability as an approximation to binomial probabilities by Abraham de Moivre. Laplace and Gauss proposed the distribution as a "law of errors" to describe the variability of measurement errors in the physical sciences.
Definition: The random variable X is said to have a normal distribution if its probability density function is given by

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty, \tag{3.5}$$

where μ and σ^2 are the parameters of the distribution. We can show that

$$E(X) = \mu \qquad Var(X) = \sigma^2.$$

We write

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

Theorem 3.2.3. *Suppose $X \sim N(\mu,\sigma^2)$. Then*

$$Z = \frac{(X-\mu)}{\sigma} \sim N(0,1)$$

the **standard normal distribution**.

Proof: We have

$$Z = \frac{X-\mu}{\sigma} \Rightarrow X = \sigma Z + \mu.$$

This implies $dx = \sigma dz$, and

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{z^2}{2}}|\sigma| = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}. \tag{3.6}$$

■
To show the function in (3.6) is a valid pdf, we need to show

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz = 1.$$

We have

$$\begin{aligned}\left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz\right]^2 &= \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz\right)\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt\right) \\ &= \frac{4}{2\pi}\int_0^{\infty}\int_0^{\infty} e^{-(z^2+t^2)/2}dzdt.\end{aligned}$$

We can make a transformation to polar coordinates. Let

$$z = r\cos\theta \qquad t = r\sin\theta.$$

Then

$$dzdt = rd\theta dr \qquad t^2 + z^2 = r^2.$$

The double integral becomes

$$\begin{aligned}\frac{4}{2\pi}\int_0^\infty\int_0^{\pi/2} e^{-r^2/2} r d\theta dr &= \frac{4}{2\pi}\frac{\pi}{2}\int_0^\infty re^{-r^2/2}dr \\ &= \int_0^\infty re^{-r^2/2}dr = -e^{-r^2/2}\Big|_0^\infty = 1.\end{aligned}$$

■

Remarks:

1. The mode (point on the horizontal axis where the curve is maximum) occurs at $x = \mu$. [Differentiate the pdf wrt x and set the derivative equal to 0.]
2. The graph of this function is a symmetric bell shaped curve, with the point of symmetry being μ.
3. The points of inflection occur at $x = \mu \pm \sigma$. The inflection points are where the curve changes from concave to convex.
4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.

Theorem 3.2.4. *If $X \sim N(\mu, \sigma^2)$, we have*

$$E(X) = \mu \qquad\qquad Var(X) = \sigma^2. \tag{3.7}$$

■

Definition: If $X \sim N(\mu, \sigma^2)$, the cumulative distribution function is given by

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt. \tag{3.8}$$

3.2.4 Gamma Distribution

The Gamma distribution has been widely used in the reliability and survival literature to model the lifetime of mechanical or biological systems.

Definition: The random variable X is said to have a Gamma Distribution if its pdf is given by

$$f(x|\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}, & x > 0; \\ 0, & \text{otherwise.}\end{cases}$$

where $\alpha, \beta > 0$. The parameter α is called the shape parameter (influences peakedness), and β is called the scale parameter (influences spread). We write $X \sim G(\alpha, \beta)$. Here $\Gamma(\alpha)$ is the **gamma function** defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$$

Integrating the function by parts, we get

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1).$$

When $\alpha = n$, a positive integer, repeated applications of the above result yield the following:

$$\Gamma(n) = (n-1)!$$

Remark: We can use the normal integral to show that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{3.9}$$

We have shown that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1 \Rightarrow \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \frac{1}{2}.$$

Let

$$t = z^2/2 \Rightarrow z^2 = 2t \qquad \Rightarrow dt = zdz.$$

We have

$$\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2t}} e^{-t} dt = \frac{1}{2}$$
$$\Rightarrow \Gamma\left(\frac{1}{2}\right) = \int_{0}^{\infty} t^{-1/2} e^{-t} dt = \frac{2\sqrt{\pi}}{2} = \sqrt{\pi}.$$

■

Theorem 3.2.5. *If $X \sim G(\alpha, \beta)$, we have*

$$\mu = \alpha\beta \qquad \sigma^2 = \alpha\beta^2. \tag{3.10}$$

■

Remarks:

1. When $\alpha = 1$, we get the exponential distribution with pdf

$$f(x|\beta) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0; \\ 0, & \text{otherwise.} \end{cases}$$

   where $\beta > 0$. We write $X \sim E(\beta)$.

2. If $\alpha = p/2$, where p is an integer and $\beta = 2$, then X is said to have the ***chi squared distribution with** p **degrees of freedom***. We write $X \sim \chi^2(p)$.

The Lack of Memory Property of the Exponential

The exponential distribution has the lack of memory property. In other words, regardless of the age of the product, there is no wearing out and the product is "as good as new". Mathematically, for $s > t \geq 0$,

$$P(X > s|X > t) = P(X > s - t). \tag{3.11}$$

One can also show that the converse is true, i.e. if a pdf satisfies (3.43), it must be the exponential pdf. From (3.43), we have

$$P(X > s|X > t) = \frac{P(X > s)}{P(X > t)} = P(X > s - t).$$

Therefore

$$\begin{aligned}\frac{1-F_X(s)}{1-F_X(t)} &= 1-F_X(s-t) \\ \Rightarrow \bar{F}_X(s) &= \bar{F}_X(s-t)\bar{F}_X(t).\end{aligned}$$

This is Cauchy's functional equation and under certain regularity conditions has the solution

$$\bar{F}(x) = e^{-cx}.$$

The exponential distribution is the only distribution with the lack of memory property. This is called a characterization of the distribution.

Relationship to the Poisson Process

If events follow a Poisson process with rate λ (the average number of events per unit time), and if T represents the waiting time from any starting point until the occurrence of the next event, then $T \sim E(1/\lambda)$. We have

$$\begin{aligned}P(T > t) &= P(\text{no occurrences in interval of length } t) \\ &= 1 - e^{-\lambda t}.\end{aligned}$$

Let T_1 represent the waiting time until the first event, T_2 the waiting time between the occurrence of the first and second event etc. Then $T_1 + \ldots T_r$ has a Gamma distribution with parameter $\alpha = r$ and $\beta = 1/\lambda$, i.e. the waiting time until r events have occurred is distributed as a Gamma random variable.

Theorem 3.2.6. *If $X \sim E(\beta)$, then $Y = X^{\frac{1}{\gamma}}$ has a Weibull distribution with parameters γ and β, with pdf given by*

$$f_Y(y|\gamma,\beta) = \frac{\gamma}{\beta} y^{\gamma-1} e^{-y^{\gamma}/\beta}, \qquad y > 0; \gamma, \beta > 0. \tag{3.12}$$

■

Applications of the Gamma and Exponential Distributions

Example: The response time X at a certain on-line computer terminal (the elapsed time between the end of a user's inquiry and the beginning of the system's response to that inquiry) has an exponential distribution with average response time equal to 5 seconds.

(a) Find the probability that the response time is at most 10 seconds.

(b) Find the probability that the response time exceeds 15 seconds.

Example: Assume that arrival times at a drive-through window follow a Poisson process with an average rate $\lambda = 0.2$ arrivals per minute. Find the probability that the third customer arrives within 20 minutes of opening.

3.2.5 Beta Distribution

Definition: The random variable X is said to have a Beta Distribution if its pdf is given by

$$f(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 < x < 1 \tag{3.13}$$

where $\alpha, \beta > 0$. We write $X \sim Beta(\alpha,\beta)$. Here $B(\alpha,\beta)$ is the **beta function** defined by

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \tag{3.14}$$

Remarks:

1. Since the range of X is the unit interval, the Beta distribution is used to model proportions.
2. If $\alpha = \beta = 1$, we get the Uniform distribution.

Theorem 3.2.7. *If $X \sim Beta(\alpha,\beta)$, we have*

$$E(X^n) = \frac{B(n+\alpha,\beta)}{B(\alpha,\beta)} \qquad n > -\alpha. \tag{3.15}$$

Proof: We have

$$\begin{aligned} E(X^n) &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^n x^{\alpha-1}(1-x)^{\beta-1} dx \\ &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{n+\alpha-1}(1-x)^{\beta-1} dx. \end{aligned}$$

The integral is the kernel of a $B(n+\alpha,\beta)$ random variable provided

$$n + \alpha > 0 \Rightarrow n > -\alpha.$$

Therefore

$$E(X^n) = \frac{B(n+\alpha,\beta)}{B(\alpha,\beta)}.$$

■

Theorem 3.2.8. *If $X \sim Beta(\alpha,\beta)$, we have*

$$E(X) = \frac{\alpha}{\alpha+\beta} \qquad V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{3.16}$$

■

3.2.6 Cauchy Distribution

Definition: The random variable X is said to have a Cauchy Distribution if its pdf is given by

$$f(x|\theta) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2} \qquad -\infty < x < \infty \tag{3.17}$$

where $\theta \in \mathcal{R}$. We write $X \sim C(\theta)$.

We have

$$\int_{-\infty}^{\infty} \frac{1}{\pi} \frac{1}{1+(x-\theta)^2} = \frac{1}{\pi} \arctan(x-\theta)|_{-\infty}^{\infty} = 1.$$

Remarks:

1. The Cauchy is a symmetric bell shaped distribution.
2. It has thicker tails than the normal.

Theorem 3.2.9. *Let $X \sim C(0)$. The moments of order < 1 exists, but moments of order ≥ 1 do not exist.*

Proof: We have

$$E[|X|^{\alpha}] = \frac{2}{\pi} \int_0^{\infty} \frac{x^{\alpha}}{1+x^2} dx.$$

The integral converges if $\alpha < 1$ and diverges if $\alpha \geq 1$.
In particular, the mean does not exist (See Page 26, Chapter 2). The mgf does not exist.
■

Theorem 3.2.10. *Let $X \sim C(\theta)$. The median is given by θ.*

Proof: The median of an absolutely continuous distribution is defined as the value x for which

$$P(X \leq x) = 0.5.$$

For the Cauchy distribution, we have

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{1}{\pi} \frac{1}{1+(x-\theta)^2} \\ &= \frac{1}{\pi} \arctan(t-\theta)|_{-\infty}^{x} \\ &= \frac{1}{2} + \frac{1}{\pi} \arctan(x-\theta). \end{aligned} \tag{3.18}$$

Setting $F_X(x) = 0.5$, we have

$$x = \theta.$$

■

3.2.7 Lognormal Distribution

Definition: If X is a random variable that is positive, and $Y = \log X$ is normally distributed, then X is said to have a Lognormal distribution with pdf given by

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}, \qquad x > 0, \tag{3.19}$$

where $\mu \in \mathcal{R}$ and $\sigma > 0$.

Remarks:

1. To show that the pdf integrates to 1, use the transformation $Y = \log X$.

2. We have

$$E(X) = e^{\mu+(\sigma^2/2)},$$

and

$$Var(X) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}.$$

These expressions can be obtained by using the relationship to the normal distribution.

3. The lognormal is a right skewed distribution that resembles the Gamma.

3.2.8 Laplace or Double Exponential Distribution

Definition: The random variable X is said to have the double exponential distribution if its pdf is given by

$$f(x|\mu,\sigma) = \frac{1}{2\sigma}e^{-\frac{|x-\mu|}{\sigma}} \qquad x \in \mathcal{R}, \tag{3.20}$$

where $\mu \in \mathcal{R}$ and $\sigma > 0$.

Remarks:

1. The double exponential is a symmetric distribution with heavier tails than the normal. It is not bell shaped.
2. We have

$$E(X) = \mu \qquad Var(X) = 2\sigma^2.$$

3. It is not differentiable at $x = \mu$.

3.2.9 Probability Integral Transformation

Theorem 3.2.11. *Let X have a continuous cdf $F_X(x)$ and let $Y = F_X(X)$. Then $Y \sim U(0,1)$, i.e.*

$$P[Y \le y] = y \qquad 0 < y < 1.$$

Proof: If $F_X(x)$ is strictly increasing, the inverse F_X^{-1} is well defined by

$$F_X^{-1}(y) = x \Leftrightarrow F_X(x) = y.$$

We have

$$\begin{aligned} P(Y \le y) &= P[F_X(X) \le y] \\ &= P[X \le F_X^{-1}(y)] \\ &= F_X[F_X^{-1}(y)] = y \qquad 0 < y < 1. \end{aligned}$$

If $F_X(x)$ is not strictly increasing, i.e. it is constant over some interval, then F_X^{-1} is not well defined. In the graph, any x satisfying $x_1 \le x \le x_2$ satisfies

$$F_X(x) = y.$$

Define

$$\begin{aligned} &F_X^{-1}(y) = \inf\{x : F_X(x) \ge y\} = x_1 \\ &F_X^{-1}(1) = \infty \ \text{ if } \ F(x) < 1 \ \ \forall \ \ x \\ &F_X^{-1}(0) = -\infty. \end{aligned}$$

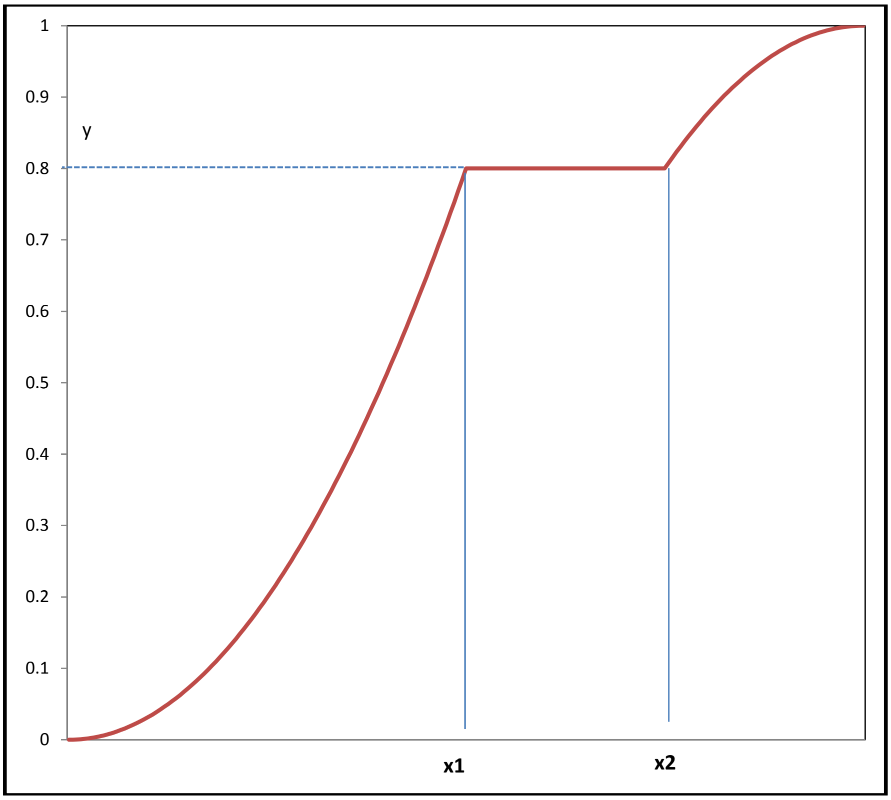

Figure 3.1: Inverse of the CDF

This definition of F_X^{-1} agrees with the usual inverse if $F_X(.)$ is strictly increasing. The proof of the theorem is then as follow:

$$\begin{aligned} P[Y \leq y] &= P[F_X(X) \leq y] \\ &= P[F_X^{-1}[F_X(X)] \leq F_X^{-1}(y)] \end{aligned}$$

since F_X^{-1} is an increasing function.
If F_X is strictly increasing, $F_X^{-1}[F_X(x)] = x$. If F_X is flat, then $F_X^{-1}[F_X(x)]$ may not equal x. For $x \in [x_1, x_2]$, $F_X^{-1}[F_X(x)] = x_1$.
However,

$$P[F_X^{-1}[F_X(X)] \leq F_X^{-1}(y)] = P[X \leq F_X^{-1}(y)]$$

since

$$P[X \leq x] = P[X \leq x_1] \qquad \text{for any} \ \ x \in [x_1, x_2].$$

Therefore

$$\begin{aligned} P[Y \leq y] &= P[X \leq F_X^{-1}(y)] \\ &= F_X[F_X^{-1}(y)] \\ &= y \end{aligned}$$

since F_X is continuous.
■

The probability integral transformation provides a way to generate random variables with specific distributions.

Theorem 3.2.12. *If $U \sim U(0,1)$, then $X = F_X^{-1}(U)$ has cdf $F_X(.)$.*

Proof: We have

$$\begin{aligned} F_X(x) = P(X \leq x) &= P[F_X^{-1}(U) \leq x] \\ &= P[U \leq F_X(x)] = F_X(x). \end{aligned}$$

■

We generate a uniform random number between 0 and 1 and solve for x in

$$F_X(x) = u$$

where F_X is the required cdf.

Example: Let $U \sim U(0,1)$. We have

$$f_U(u) = \begin{cases} 1, & 0 < u < 1; \\ 0, & \text{o.w.} \end{cases}$$

Let $Y = -ln\ U$. This implies $U = e^{-Y}$. The range of Y is $(0, \infty)$. Using Theorem 3.1.2, we have

$$f_Y(y) = f_U(e^{-y})|-e^{-y}| = e^{-y}.$$

This is the pdf of the Exponential distribution.
■

Example: Let X be a continuous r.v. with pdf

$$f_X(x) = \begin{cases} \frac{1}{2}, & 1 < |x-2| < 2; \\ 0, & \text{o.w.} \end{cases}$$

The range of x is $(0,1)$ and $(3,4)$. The cdf of X is

$$F_X(x) = \begin{cases} 0, & x < 0; \\ \frac{1}{2}x, & 0 \le x < 1; \\ \frac{1}{2}, & 1 \le x \le 3; \\ \frac{1}{2} + \frac{x-3}{2}, & 3 < x < 4; \\ 1, & x \ge 4. \end{cases}$$

The cdf is not 1-1. The inverse is

$$F_X^{-1}(y) = \begin{cases} 2y, & 0 < y \le 0.5; \\ 2(y+1), & 0.5 < y < 1. \end{cases}$$

■

Example: $X \sim Bin(1,p)$.

$$F_X(x) = \begin{cases} 0, & x < 0; \\ p, & 0 \le x < 1; \\ 1, & x \ge 1. \end{cases}$$

$$F_X^{-1}(y) = \begin{cases} 0, & 0 < y \le p; \\ 1, & p < y \le 1. \end{cases}$$

■