Continuous Distributions

1 Random Variables of the Continuous Type

Density Curve

A smooth curve that fits the distribution

Probability density is not probability!

Probability density function, \( f(x) \)

Use a mathematical model to describe the variable.

Continuous Distribution

Probability Density Function (p.d.f) of a random variable \( X \) of continuous type with a space \( S \) is an integrable function, \( f(x) \), that satisfies the following conditions:

1. \( f(x) \geq 0, x \in S \),
2. \( \int_S f(x) \, dx = 1 \) (Total area under curve is 1.)
3. For \( a \) and \( b \) in \( S \), \( P(a < X < b) = \int_a^b f(x) \, dx \)

Meaning of Area Under Curve

Example: What percentage of the distribution is in between 72 and 86?

\( P(X = 72) = 0 \), density is not probability.

\( P(72 \leq X \leq 86) = P(72 < X < 86) = P(72 \leq X < 86) = P(72 < X \leq 86) \)
Continuous Distributions

Example:
If the density function of a continuous distribution is
\[ f(x) = \begin{cases} 8x, & \text{for } 0 < x < 0.5 \\ 0, & \text{elsewhere} \end{cases} \]
Find the proportion of values in this distribution that is less than 1/4.
The area under the \( f(x) \) between 0 and 1/4 is
\[ \int_0^{1/4} 8x \, dx = 8 \left[ \frac{x^2}{2} \right]_0^{1/4} = 4 \left( \frac{1}{4} \right)^2 - 4(0)^2 = 4/16 = 1/4. \]

Example:
If the density function of a continuous distribution \( X \), waiting time between arrivals of cars at a intersection, is
\[ f(x) = \frac{1}{5} e^{-\frac{x}{5}}, \quad \text{for } x > 0 \]
Find the probability that the waiting time (in seconds) till the next arrival of car at this intersection is more than 3 seconds.
The area under the \( f(x) \) and \( x > 3 \) is
\[ \int_3^{\infty} \frac{1}{5} e^{-\frac{x}{5}} \, dx = \lim_{b \to \infty} \left[ -e^{-\frac{x}{5}} \right]_3^b = \lim_{b \to \infty} (-e^{-\frac{b}{5}}) - (-e^{-\frac{3}{5}}) = 0 - (-0.5488) = 0.5488 \]

Cumulative Distribution Function
The cumulative distribution function (c.d.f. or distribution function, d.f.) of a continuous random variable is defined as
\[ F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t) \, dt \]
- \( F(-\infty) = 0, \quad F(\infty) = 1 \)
- \( P(a < X < b) = F(b) - F(a) \)
- \( F(x) = f(x) \) if derivative exists

Review of Calculus
\[ \int x^n \, dx = \frac{1}{n+1} x^{n+1} + c \]
\[ \int e^{ax} \, dx = \frac{1}{a} e^{ax} + c \]

Example:
If the density function of a continuous distribution \( X \), waiting time between arrivals of cars at a intersection, is
\[ f(x) = \frac{1}{5} e^{-\frac{x}{5}}, \quad \text{for } x > 0 \]
Find the probability that the waiting time (in seconds) till the next arrival of car at this intersection is less than 3 seconds.
The area under the \( f(x) \) below 3 is
\[ \int_0^{3} \frac{1}{5} e^{-\frac{x}{5}} \, dx = \left[ -e^{-\frac{x}{5}} \right]_0^3 = (-e^{-\frac{3}{5}}) - (-e^{0}) = e^{-\frac{3}{5}} - 1 - 0.5488 = 0.4522 \]

Example:
If the p.d.f. of a continuous distribution \( X \), waiting time between arrivals of cars at a intersection, is, where \( \theta \) is a constant parameter of this distribution.
\[ f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad \text{for } x > 0 \]
Find the distribution function.
\[ F(x) = \int_{-\infty}^{x} f(t) \, dt = \int_0^{x} \frac{1}{\theta} e^{-\frac{t}{\theta}} \, dt = \left[ -e^{-\frac{t}{\theta}} \right]_0^x = 1 - e^{-\frac{x}{\theta}} \]
\[ F(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1 - e^{-\frac{x}{\theta}}, & 0 < x < \infty \end{cases} \]
Continuous Distributions

Measure of Center for a Continuous Distribution

The mean value (expected value) of a continuous random variable (distribution) $X$, denoted by $\mu_X$ or just $\mu$ (or $E[X]$) is defined as

$$
\mu_X = \int_{-\infty}^{\infty} x \cdot f(x) \, dx
$$

Measure of Spread for a Continuous Distribution

The variance of a continuous random variable (distribution) $X$, denoted by $\sigma_X^2$ or just $\sigma^2$ (or $Var[X]$) is defined as

$$
\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) \, dx
$$

The standard deviation of $X$ is

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

Moment Generating Function for a Continuous Distribution

The moment generating function, if exists, of a continuous random variable (distribution) $X$, denoted by $M(t)$ is defined as

$$
M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx,
\quad -h < t < h
$$

Example:

If the density function of a continuous distribution is

$$
f(x) = \begin{cases} 
8x & \text{for } 0 < x < 0.5 \\
0 & \text{elsewhere}
\end{cases}
$$

Find the mean and variance of this distribution

The mean is

$$
\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \, dx = \int_{0}^{0.5} x \cdot f(x) \, dx
$$

$$
= \int_{0}^{0.5} x \cdot 8x \, dx = 8 \int_{0}^{0.5} x^2 \, dx = \frac{8}{3} \left( \frac{0.125 - 0}{3} \right) + \frac{1}{3}
$$

The Percentile

The $(100p)$th percentile (quantile of order $p$) is the number $\pi_p$, such that the area under $f(x)$ to the left of $\pi_p$ is $p$.

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

Median is the 50th percentile.
Continuous Distributions

**Median of a distribution**

**Example:**
If the density function of a continuous distribution is

\[ f(x) = \begin{cases} 
8x, & \text{for } 0 < x < 0.5 \\
0, & \text{elsewhere}
\end{cases} \]

Find the median of this distribution.

Median is \( c \) such that

\[ \int_0^c 8x \, dx = 0.5 \]

What is \( c \)?

\[ \int_0^c 8x \, dx = \frac{8x^2}{2} \bigg|_0^c = 0.5 \]

\[ \Rightarrow 4c^2 - 0 = 0.5 \]

\[ \Rightarrow c^2 = \frac{0.5}{4} \]

\[ \Rightarrow c = \sqrt{\frac{0.5}{4}} = 0.354 \]

---

**Percentile**

**Example:**
If the density function of a continuous distribution is

\[ f(x) = \begin{cases} 
8x, & \text{for } 0 < x < 0.5 \\
0, & \text{elsewhere}
\end{cases} \]

Find the 25th percentile of this distribution.

25th percentile is \( c \) such that

\[ \int_0^c 8x \, dx = 0.25 \]

What is \( c \)?

\[ \int_0^c 8x \, dx = \frac{8x^2}{2} \bigg|_0^c = 0.25 \]

\[ \Rightarrow 4c^2 - 0 = 0.25 \]

\[ \Rightarrow c^2 = \frac{0.25}{4} \]

\[ \Rightarrow c = \sqrt{\frac{0.25}{4}} = 0.25 \]

---

**Sample Quantile**

Let \( y_1 \leq y_2 \leq \ldots \leq y_n \) be the order statistics associated with the sample \( x_1, x_2, \ldots, x_n \), then \( y_r \) is called the sample quantile of order \( r/n + 1 \) as well as \( \frac{100}{n + 1} \) percentile.

**Example:** 1, 3, 7, 8, 9, 13

\( n = 6 \), and the value \( 8, y_4 \),

is the \((4/(6+1))\)th quantile of the distribution of the sample, i.e. 0.571 sample quantile or 57.1 percentile.

---

**Examine Distribution with Quantile-Quantile Plot**

Make a plot using data in previous slide.
Continuous Distributions

2 The Uniform and Exponential Distributions

Uniform Distribution
The continuous random variable $X$ has a **uniform distribution** if its p.d.f. is equal to a constant on its support. If the support is the interval $[a, b]$, then its p.d.f. is

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

It is usually denoted as $U(a, b)$.

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{elsewhere} \end{cases}$$

**Pseudo-Random Number Generator** on most computers $U(0, 1)$

Exponential Distribution
The continuous random variable $X$ has an **exponential distribution** if its p.d.f. is

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & 0 < x \\ 0, & \text{elsewhere} \end{cases}$$

where $\theta$ is the mean of the distribution.

* $X$ can be the waiting time until next success in a Poisson process.
Continuous Distributions

Exponential Distribution

**Example:** Let $X$ have an exponential distribution with a mean of 30, what is the first quartile of this distribution?

$$F(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1 - e^{-x/30}, & 0 \leq x < \infty \end{cases}$$

$$F(2.5) = 1 - e^{-2.5/30} \Rightarrow -2.5/30 = 0.75 \Rightarrow -\frac{\ln(0.75)}{\theta} = 8.63$$

$\theta = 30 \Rightarrow P(0.25) = 1 - e^{-0.25/8.63} = 0.918$

Exponential Distribution

Let $W$ be the waiting time until next success in a Poisson process in which the average number of success in unit interval is $\lambda$, then, for $w \geq 0$

$$F(w) = 1 - e^{-w/\theta} \Rightarrow \text{d.f. of exponential distribution.}$$

$$f(w) = \frac{1}{\theta} e^{-w/\theta} \Rightarrow \text{p.d.f. of exponential distribution.}$$

Exponential Distribution

Suppose that number of arrivals of customers follows a Poisson process with a mean of 10 per hour. What is the probability that the next customer will arrive within 15 minutes? (15 min. = 0.25 hour)

$$P(0.25) = 1 - e^{-0.25/0.1} = 0.918$$

Exponential Distribution

Let $W$ be the waiting time until next success in a Poisson process in which the average number of success in unit interval is $\lambda$, then, for $w \geq 0$

$$F(w) = 1 - e^{-w/\theta} \Rightarrow P(0.25) = 1 - e^{-0.25/0.1} \Rightarrow \text{p.d.f. of exponential distribution.}$$

Gamma Distribution

The continuous random variable $X$ has a Gamma distribution, $\Gamma(\alpha, \theta)$, if its p.d.f. is

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}, & \text{for } 0 \leq x \\ 0, & \text{elsewhere.} \end{cases}$$

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

* $X$ can be the waiting time until $\alpha$-th success in a Poisson process.
Continuous Distributions

Gamma Distribution

The mean, variance, and m.g.f. of a continuous random variable $X$ that has an *Gamma distribution* are:

$$
\mu = \alpha \theta, \quad \sigma^2 = \alpha \theta^2, \\
M(t) = \frac{1}{(1 - \theta t)^\alpha}, \quad t < \frac{1}{\theta}
$$

- Gamma(1, $\theta$) $\Rightarrow$ Exponential Distribution
- Gamma(2, $r/2$) $\Rightarrow$ Ch-square with d.f. = $r$.

Special Notation $\chi^2_{\alpha}(r)$

Let $X$ be a random that has Chi-square distribution with degrees of freedom $r$.

$$
P[X \geq \chi^2_{\alpha}(r)] = \alpha
$$

Example: Find $\chi^2_{0.05}(3) = 7.815$

Try This!

- Find $\chi^2_{0.1}(7) = ?$
- Find $\chi^2_{0.025}(4) = ?$
- Find the 10th percentile from a $\chi^2$ distribution with degrees of freedom 6.
Continuous Distributions

A Good Generator

\[ X_i = (397,204,094 X_{i-1}) \mod (2^{31} - 1) \]

• Simple
• Widespread use
• Long cycle length \(2^{31} - 2\) (all numbers besides 1 and \(2^{31} - 1\) can be generated.)
• \(U_i = X_i / (2^{31} - 1),\ U_i \sim U(0,1)\)

Methods for Generating Non-Uniform RN’s

• CDF Inversion
• Transformations
• Accept/Reject Methods
• ...

CDF Inversion

**Theorem:** Let \( U \) have a distribution that is \( U(0,1) \). Let \( F(x) \) have the properties of a distribution function of the continuous type with \( F(a) = 0 \) and \( F(b) = 1 \), and suppose that \( F(x) \) is strictly increasing on the support \( a < x < b \), where \( a \) and \( b \) could be \(-\infty\) and \( \infty \), respectively. Then the random variable \( X \) defined by \( X = F^{-1}(U) \) is a continuous random variable with distribution function \( F(x) \).

**Proof:**

Let \( X = F^{-1}(U) \), the distribution function of \( X \) is \( P(X \leq x) = P[U \leq F(x)] \), \( a < x < b \).

Since \( F(x) \) is strictly increasing, \( \{F^{-1}(U) \leq x\} \) is equivalent to \( \{U \leq F(x)\} \) and hence

\[ P(X \leq x) = P[U \leq F(x)], \quad a < x < b. \]

But \( U \) is \( U(0,1) \); so \( P(U \leq u) = u \) for \( 0 < u < 1 \), and accordingly,

\[ P(X \leq x) = P[U \leq F(x)] = F(x), \quad 0 \leq F(x) < 1. \]

That is the distribution function of \( X \) is \( F(x) \).

Generate random numbers from Exponential distribution, \( \theta = 10 \)

• \( f(x) = 1 - e^{-x/10} \) and \( x = F^{-1}(y) = -10 \cdot \ln(1 - y) \)
• Use uniform \( U(0,1) \) random number generator to generate random numbers \( y_1, y_2, ..., y_n \).
• The exponentially distributed random numbers \( x_i \)'s would be \( x_i = -10 \cdot \ln(1 - y_i), \quad i = 1, ..., n \).
• Therefore, if the uniform random number generator generates a number \( 0.1514 \), then \( 1.6417 = -10 \cdot \ln(1 - 0.1514) \) would be a random observation from exponential distribution with \( \theta = 10 \).

Try this!!!

An \( U(0,1) \) random number generator has generated a value of 0.26.

1. Use the CDF Inversion method to convert this \( U(0,1) \) random number to simulate an observation from an exponential distribution with \( \theta = 12 \).

\[ x_i = F^{-1}(y_i) = -12 \cdot \ln(1 - y_i) \]

2. Use CDF Inversion method to generate a random number from the continuous random variable that has the following p.d.f.

\[ f(x) = \begin{cases} 8x & , \text{ for } 0 < x < 0.5 \\ 0 & , \text{ elsewhere} \end{cases} \]