Continuous Distributions

Probability Density Function –

Definition:
1. \( f(x) \geq 0 \).
2. \( \int_{-\infty}^{\infty} f(x) \, dx = 1 \).
3. \( P(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx \).

You use \( f(x) \) to calculate an area that represents the probability that \( X \) takes on a value in \([a,b]\).
Example:
Determine the probability that $X$ assumes a value between 0.6 and 1.2.

Where $f(x) = \begin{cases} x & \text{for } 0 < x < 1 \\ 2-x & \text{for } 1 \leq x < 2 \\ 0 & \text{elsewhere} \end{cases}$

*For a continuous random variable, $P(X = x) = 0$.

Cumulative distribution function:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \quad \text{for } -\infty < x < \infty.$$ 

$$P(a \leq X \leq b) = F(b) - F(a)$$
Example:
Suppose the cumulative distribution function of the random variable X is:

\[ F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
0.2x & \text{if } 0 \leq x < 5 \\
1 & \text{if } x \geq 5 
\end{cases} \]

Determine: The probability density function of x and P(x < 2.8).

Mean and Variance of a Continuous Random Variable:

\[ \mu = \text{mean} = E(x) = \int_{-\infty}^{\infty} xf(x) \, dx \]

\[ \text{variance} = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) \, dx \]
Uniform Distribution:

\[ f(x) = \frac{1}{b-a} \quad a \leq x \leq b \]

\[ F(x) = \int_{-\infty}^{x} f(u) du = \frac{1}{b-a} \int_{a}^{x} du = \frac{x-a}{b-a} \quad a \leq x \leq b \]

\[ E(x) = \mu = \int_{-\infty}^{\infty} xf(x) dx = \int_{a}^{b} x \left( \frac{1}{b-a} \right) dx = \frac{b+a}{2} \]

\[ \sigma^2 = \frac{1}{12} (b-a)^2 \]

Example:
A bomb is to be dropped along a mile long line that stretches across a practice target. The target center is at the midpoint of the line. The target will be destroyed if the bomb falls within 0.1 miles to either side of the center. Find the probability that the target is destroyed if the bomb falls randomly along the line.
Exponential Distribution:

\[ f(x) = \lambda e^{-\lambda x} \quad \text{where} \quad \lambda = \text{average rate of occurrence} \]

\[ F(x) = 1 - e^{-\lambda x} \]

\[ E(x) = \mu = \frac{1}{\lambda} \]

\[ \sigma^2 = \frac{1}{\lambda^2} \]

*Events are Poisson distributed, time between events will be exponentially distributed, thus the derivation of this model.*

Example:
At a stop sign location on a cross street, vehicles require headways of 6 seconds or more in the main street traffic before being able to cross. If the total flow rate of the main street traffic is 1200 vph, what is the probability that any given headway will be greater than 6 seconds?

Example:
An electronic component is known to have a useful life represented by an exponential density with failure rate of $10^{-5}$ failures per hour. What fraction of the components will fail before the mean life?
The exponential distribution is unique in that it can be said to be “memoryless.” This means that the probability of a success in a certain time period does not change if the start time of the interval changes.

Example: The lifetime of a particular integrated circuit has an exponential distribution with mean 2 years. Find the probability that the circuit lasts longer than three years:

Now, suppose the circuit is now four years old and is still functioning. Find the probability that it functions for more than three additional years. Compare this with the previous probability (a new circuit functions for more than three years). So, we are interested in a conditional probability:

\[
\therefore \text{After 4 years, the probability of the circuit lasting more than three years is the same as if it is new, thus the lack of memory property.}
\]

**Gamma Distribution:**
The exponential distribution is a special case of the Gamma distribution, where X is the interval length until r counts occur in a Poisson process.
\[ f(x) = \begin{cases} \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} & x > 0 \\ \end{cases} \]

**Gamma Function:** \[ \Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx \]

\[ F(x) = P(T \leq x) = \begin{cases} 1 - \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} & x > 0 \\ 0 & \text{elsewhere} \end{cases} \]

\[ E(x) = \mu = \frac{r}{\lambda} \]

\[ \sigma^2 = \nu(x) = \frac{r}{\lambda^2} \]

Figure 4-36 (Montgomery text)
The Gamma Function: \( \Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx \)

Properties:
1. \( \Gamma(1) = 1 \)
   
   For \( r>1 \):
2. \( \Gamma(r) = (r-1) \Gamma(r-1) \)
3. \( \Gamma(\frac{1}{2}) = \sqrt{\pi} \)
4. \( \Gamma(r+1) = r! \)
   
   *for large values of \( r \), \( r! \) may be approximated using Stirling’s approximation.

Stirling’s Approximation to \( n! \):
\[
 n! \cong \sqrt{2\pi n} \left( \frac{n}{e} \right)^n \quad n \to \infty
\]

Accurate to within 1\% for \( n>10 \) and within 0.1\% for \( n>100 \).

Example:
Suppose the survival time, in weeks, of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with \( r = 8 \) and \( \lambda = 1/15 \). Find:
   
   a.) The expected survival time;
   
   b.) variance;
   
   c.) the probability that a mouse survives between 60 and 120 weeks;
   
   d.) the probability that a mouse survives at least 30 weeks.
*GAMMA FUNCTION*

Values of $\Gamma(n) = \int_0^\infty e^{-x}x^{n-1}dx; \Gamma(n+1) = n\Gamma(n)$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Gamma(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.000000</td>
</tr>
<tr>
<td>1.01</td>
<td>0.99433</td>
</tr>
<tr>
<td>1.02</td>
<td>0.98864</td>
</tr>
<tr>
<td>1.03</td>
<td>0.98355</td>
</tr>
<tr>
<td>1.04</td>
<td>0.97844</td>
</tr>
<tr>
<td>1.05</td>
<td>0.97350</td>
</tr>
<tr>
<td>1.06</td>
<td>0.96874</td>
</tr>
<tr>
<td>1.07</td>
<td>0.96415</td>
</tr>
<tr>
<td>1.08</td>
<td>0.95973</td>
</tr>
<tr>
<td>1.09</td>
<td>0.95546</td>
</tr>
<tr>
<td>1.10</td>
<td>0.95135</td>
</tr>
<tr>
<td>1.11</td>
<td>0.94739</td>
</tr>
<tr>
<td>1.12</td>
<td>0.94359</td>
</tr>
<tr>
<td>1.13</td>
<td>0.93993</td>
</tr>
<tr>
<td>1.14</td>
<td>0.93642</td>
</tr>
<tr>
<td>1.15</td>
<td>0.93304</td>
</tr>
<tr>
<td>1.16</td>
<td>0.92980</td>
</tr>
<tr>
<td>1.17</td>
<td>0.92670</td>
</tr>
<tr>
<td>1.18</td>
<td>0.92373</td>
</tr>
<tr>
<td>1.19</td>
<td>0.92088</td>
</tr>
<tr>
<td>1.20</td>
<td>0.91817</td>
</tr>
<tr>
<td>1.21</td>
<td>0.91558</td>
</tr>
<tr>
<td>1.22</td>
<td>0.91311</td>
</tr>
<tr>
<td>1.23</td>
<td>0.91075</td>
</tr>
<tr>
<td>1.24</td>
<td>0.90852</td>
</tr>
</tbody>
</table>

*For large positive values of $x$, $\Gamma(x)$ approximates Stirling's asymptotic series*

$$x^x e^{-x} \sqrt{\frac{2\pi}{x}} \left[ 1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320x^4} + \cdots \right].$$
The Incomplete Gamma Function:

The incomplete gamma function is often used for ease of application. It is a transformation of the gamma function.

We will re-write the Gamma function:

\[ r = \alpha \quad \text{and} \quad \lambda = \frac{1}{\beta} \]

Where \( \alpha, \beta \) are parameters which determine the shape of the curve.

So, \( f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \) for \( x > 0 \).

Now, \( E(x) = \alpha \beta \) and \( V(x) = \alpha \beta^2 \).

The incomplete gamma function may then be defined as:

\[ F(x, \alpha) = \int_0^x \frac{x^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy \quad x > 0. \]

If \( X \) has a gamma distribution with parameters \( \alpha \) and \( \beta \), then for any \( x > 0 \):

\[ P(X \leq x) = F\left( \frac{x}{\beta}; \alpha \right) \]

Where \( F(\cdot; \alpha) \) is the incomplete gamma function.
Example:

Suppose the survival time, in weeks, of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 15$. Find:

a.) The expected survival time;

b.) variance;

c.) the probability that a mouse survives between 60 and 120 weeks;

d.) the probability that a mouse survives at least 30 weeks.
**Weibull Distribution:**

Used to model time until failure of various physical systems.

\[
f(x) = \frac{\beta}{\delta} \left( \frac{x}{\delta} \right)^{\beta-1} e^{-\left( \frac{x}{\delta} \right)^{\beta}} \text{ for } x > 0,
\]

where \( \delta = \text{scale parameter} \) and \( \beta = \text{shape parameter} \).

\[
F(x) = 1 - e^{-\left( \frac{x}{\delta} \right)^{\beta}}
\]

\[
E(x) = \mu = \delta \Gamma \left( 1 + \frac{1}{\beta} \right)
\]

\[
V(x) = \sigma^2 = \delta^2 \Gamma \left( 1 + \frac{2}{\beta} \right) - \delta^2 \left[ \Gamma \left( 1 + \frac{1}{\beta} \right) \right]^2
\]

---

Figure 4-27 Montgomery text
Example: (Navidi text)
Researchers suggest using a Weibull distribution to model the duration of a bake step in the manufacture of a semiconductor. Let $T$ represent the duration in hours of the bake step for a randomly chosen lot. If $T$ follows a Weibull distribution having $\beta = 0.3$, and $\delta = 10$, what is the probability that the bake step takes longer than four hours? What is the probability that it takes between two and seven hours?

The Normal Distribution:

Most widely used model for distribution of random variables. May also be called Gaussian distribution.

You can model random variables with different means and variances with the normal distribution, as long as you choose appropriate center and width of curve (mean and standard deviation).
Probability Density Function:

\[ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \]

\[ E(x) = \mu \]

\[ V(x) = \sigma^2 \]

The distribution is symmetrical, \( \therefore P(X>\mu) = P(X<\mu) = 0.5 \)

For any normal random variable:

\[ P(\mu - \sigma < X < \mu + \sigma) = 0.6827 \]

\[ P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9545 \]

\[ P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9973 \]
A normal random variable is called a standard normal random variable (z) if \( \mu = 0 \) and \( \sigma^2 = 1 \).

Cumulative Distribution Function: \( F(z) = P(Z \leq z) \)

*Tables are used to determine the value of \( F(z) \). (Table II in Appendix A.)

Ex. (from book)

\[
\begin{align*}
P(Z > 1.26) &= 1 - P(Z \leq 1.26) = 1 - 0.89616 = 0.10384 \\
P(Z < -0.86) &= 0.19490 \\
P(Z > -1.37) &= P(Z < 1.37) = 0.91465 \quad \text{(by symmetry)} \\
P(-1.25 < Z < 0.37) &= P(Z < 0.37) - P(Z < -1.25) = 0.66431 - 0.10565 = 0.53866
\end{align*}
\]
A transformation must be applied for normal distributions which do not correspond to the standard normal distribution in order to make use of the normal tables for $F(z)$.

**Z transformation:**

If $x$ is a normal random variable with $E(x) = \mu$ and $V(x) = \sigma^2$, then

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

where the transformed variable $Z$ is a standard normal random variable.

**Example:**
In diaphragms of rats, tissue respiration rate under standard temperature conditions is normally distributed with $\mu = 2.03$ and $\sigma = 0.44$.

a. What is the probability that a randomly selected rat has rate $X > 2.5$?

b. What is the probability that $X$ falls outside the interval (1.59, 2.47)?
Example:
In an industrial process, the diameter of a ball bearing is an important component part. The buyer sets specifications on the diameter to be $3.0 \pm 0.01$ cm. The implication is that no part falling outside these specifications will be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with mean 3.0 and standard deviation 0.005. On the average, how many manufactured ball bearings will be scrapped?

Example:
Gauges are used to reject all components where a certain dimension is not within the specification $1.5 \pm d$. It is known that this measurement is normally distributed with mean 1.5 and standard deviation 0.2. Determine the value $d$ such the specifications “cover” 95% of the measurements.
The Normal Approximation to the Binomial:

*Valid only for cases where n is large relative to p.  (np > 5)

If X is a binomial random variable,

\[ Z = \frac{X - np}{\sqrt{np(1 - p)}} = \frac{X - np}{\sqrt{npq}} \] is approximately a standard normal random variable.

Example:
In 6000 tosses of a die, what is the probability of obtaining a “3” between 980 and 1030 times, inclusive?

Example:
The probability that a patient recovers from a rare blood disease is 0.4. If 100 people are known to have contracted this disease, what is the probability that 30 or less survive?
Continuity correction for the binomial approximation:

Because the normal distribution is continuous and the binomial distribution is discrete, we have to do a continuity correction for $X$ in our approximation of the binomial with the normal distribution.
The Central Limit Theorem:

General definition: If a random variable $Y$ is the sum of $n$ independent random variables that satisfy certain general conditions, then for sufficiently large $n$, $Y$ is approximately normally distributed.

Central Limit Theorem – If $\bar{x}$ is the mean of a random sample of size $n$ taken from a population with mean $\mu$ and finite variance $\sigma^2$, then the limiting form of the distribution of $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ as $n \to \infty$, is the standard normal distribution $N(Z;0,1)$.

$\mu_\bar{x} = \mu$

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$

The normal approximation for $\bar{x}$ will generally be good for $n \geq 30$. It may still work for $n < 30$, but would have to be well behaved.

Example:
An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

Example:
When a batch of a certain chemical product is prepared, the amount of a particular impurity in the batch is a random variable with mean value 4.0 g and standard deviation 1.5 g. If 50 batches are independently prepared, what is the approximate probability that the sample average amount of impurity is between 3.5 and 3.8 g?