**[Normal distribution]**

Q\_1: What is a normal distribution?

A\_1: The normal distribution, one of the most representative continuous probability distributions, is also called the Gaussian distribution since Gauss demonstrated that the probability distribution of errors matches the normal distribution. Its parameters are the mean and variance.

Q\_2: Where is the normal distribution used?

A\_2: The normal distribution is frequently used because many natural phenomena and datasets tend to cluster around the mean. For example, human height, weight, and test scores mostly gather near the average, with extreme values (very tall or very short) being rare, which often leads to following a normal distribution.

Q\_3: What is variance?

A\_3: Variance is a statistical measure that indicates the degree of dispersion in the data. It is calculated by squaring the difference between each data point and the mean, and then averaging these squared values. For random variables, the variance represents the variability of the random variable in a probability distribution. It is the expected value (mean) of the squared difference between the random variable and its expected value (mean). Since variance involves squaring, its units differ from the original data. To express variability in the same units as the original data, the square root of the variance, known as the standard deviation, is used. If two normal distributions have the same variance but different means, their curves will have the same shape, but their centers will differ as determined by the mean.

Q\_4: You mentioned using the standard deviation, which is the square root of the variance, to express variability in the same units. Why is that?

A\_4: Variance shows the spread of values, but since it is calculated by squaring the differences, its units become larger than the original ones. For example, if the values are in centimeters (cm), the variance will be in square centimeters (cm²). To bring it back to the original unit (cm), we take the square root, which gives us the standard deviation. This allows us to express variability in the same units as the original data.

Q\_5: Characteristics of the normal distribution?

A\_5: 1. Symmetry: The normal distribution curve is symmetric about the mean. It has a bell shape that is symmetrical on both sides of the mean, indicating that most frequencies are concentrated near the mean, and frequencies rapidly decrease as they move away from the mean.

2. Coincidence of Mean, Median, and Mode: In a normal distribution, the mean, median, and mode all have the same value, which is the center of the distribution.

3. Shape of the Curve: The normal distribution curve is bell-shaped. It is highest at the center and gradually decreases as it extends outwards. The curve extends infinitely but never reaches zero.

4. Probability Density Function: The probability density function (PDF) of the normal distribution is given by: $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\left(x-\mu\right)^{2}/2\sigma^{2}}$

5. Standard Deviation and Curve Width: The width of the curve is determined by the size of the standard deviation. A larger standard deviation results in a wider curve, while a smaller standard deviation produces a narrower curve. The total area under the normal distribution curve is 1.

6. Probability Rule of Distribution: The probabilities associated with specific ranges in the normal distribution are as follows:

About 68.3% of the data falls within the range of the mean $\pm$ 1$\sigma$.

About 94.5% of the data falls within the range of the mean $\pm$ 2$\sigma$.

About 99.7% of the data falls within the range of the mean $\pm$ 3$\sigma$.

Q\_6: Why does the normal distribution have a bell shape?

A\_6: The normal distribution has a bell shape because it is "symmetrically distributed around the mean." Most values are clustered "near the mean," and as you move further away from the mean, the values decrease, which causes the graph to taper off on both sides.

Q\_7: Why are the mean, median, and mode the same in a normal distribution?

A\_7: The normal distribution is perfectly symmetrical. Therefore, the mean (the average of all values), the median (the middle value), and the mode (the most frequently occurring value) all lie in the center of the graph. In other words, these three values are the same because the data is symmetrically distributed.

Q\_8: What does it mean that the frequency is concentrated around the mean?

A\_8: In a normal distribution, it means that "most values" are distributed close to the mean. In other words, values frequently occur near the mean, and become less common as you move further away from the mean. This is why the normal distribution graph is highest in the middle and decreases towards the ends.

Q\_9: The normal distribution graph is bell-shaped, but why does it never actually reach 0?

A\_9: The normal distribution graph extends infinitely but never actually reaches 0. This means that there is a theoretical possibility of very extreme values occurring. However, the probability of these extreme values is so small that while the graph continues to taper off endlessly, it never actually reaches zero.

Q\_10: Why does the size of the standard deviation determine the width of the curve?

A\_10: A larger standard deviation indicates that the data is more spread out from the mean. As a result, the normal distribution graph becomes wider. Conversely, a smaller standard deviation means the data is more clustered around the mean, causing the graph to become narrower.

Q\_11: Why is the area under the normal distribution curve equal to 1?

A\_11: The area under the normal distribution curve is equal to 1 because the curve represents probabilities. The total probability of all possible outcomes must sum to 100%, or 1. Therefore, the area under the normal distribution curve, which represents the total probability, is 1.

Q\_12: What is a standard normal distribution?

A\_12: A standard normal distribution is a normal distribution with a mean of 0 and a variance of 1. Its probability density function is given by: $\phi\left(x\right)=\frac{1}{\sqrt{2\pi}}e^{-x^{2}/2}$.

The standard normal distribution is denoted by Z, and a random variable Z following this distribution is expressed as $Z\sim N\left(0,1\right)$.

Q\_13: What are the characteristics of a standard normal distribution?

A\_13: 1. Symmetry: The standard normal distribution is symmetric around the mean z=0.

2. Maximum Value: The probability density function has its maximum value at z=0.

3. Area Under the Curve: The total area under the curve is 1, which signifies that the sum of all probabilities is 1.

4. Probability Distribution Rules:

The probability that z falls between -1 and 1 is approximately 68.27%.

The probability that z falls between -2 and 2 is approximately 95.45%.

The probability that z falls between -3 and 3 is approximately 99.73%.

All normal distributions can be transformed into the standard normal distribution through a process called standardization: $Z=\frac{X-\mu}{\sigma}$

where Z follows the standard normal distribution. The standard normal distribution is used as a benchmark in statistical analysis to standardize data. This allows for comparison of datasets with different means and standard deviations.

Q\_14: The standard normal distribution is denoted by Z. Can it be represented by other symbols?

A\_14: Z is a commonly used symbol for the standard normal distribution, but other symbols can be used as well. However, in statistics, Z is the standard notation and is most commonly used.

Q\_15: What does it mean that as the sample size increases, the distribution of the sample mean approaches a standard normal distribution?

A\_15: As the sample size increases, the statement that "the mean of the sample approaches a normal distribution" means that with a larger number of data points, the distribution of the sample mean will approximate a normal distribution. This concept is known as the Central Limit Theorem. It states that regardless of the original distribution of the data, the distribution of the sample mean will tend to be normal if the sample size is sufficiently large.

Q\_16: Why is it useful to standardize all normal distributions?

A\_16: Standardizing data changes the mean to 0 and the standard deviation to 1, making it easier to compare different datasets. By converting various normal distributions to a common standard, they can be analyzed using the same reference, which facilitates comparison and analysis.

Q\_17: What is the Central Limit Theorem?

A\_17: The Central Limit Theorem states that if the sample size is sufficiently large, the distribution of the sample mean will approximately follow a normal distribution, regardless of the original distribution of the data. In other words, the mean of a large number of independent and identically distributed random variables will tend to be normally distributed. As the sample size increases, the distribution of the standardized sample mean approaches a standard normal distribution $N\left(0,1\right)$. For the theorem to apply, the random variables in the sample must be independent, come from the same distribution, and the population must have a finite variance. If the population has extremely heavy tails or infinite variance, the Central Limit Theorem may not apply, or a very large sample may be needed to achieve normality.

Q\_18: What happens if the random variables in the sample are not independent in the Central Limit Theorem?

A\_18: If the random variables in the sample are not independent, the Central Limit Theorem does not apply. This means that the sample mean may not follow a normal distribution, making it difficult to use normal distribution-based predictions and analyses. In such cases, alternative methods or additional considerations may be needed to accurately describe the distribution of the sample mean.

Q\_19: How do normal distributions differ from other probability distributions?

A\_19: The normal distribution is characterized by values being concentrated around the mean with a symmetric, bell-shaped curve. In contrast:

- Binomial Distribution: Deals with two possible outcomes (success or failure) and models the number of successes in a fixed number of trials.

- Exponential Distribution: Models the time between successive events in a process where events occur continuously and independently at a constant rate.

- Weibull Distribution: Describes situations where the failure rate changes over time, making it useful for analyzing reliability and life data where the hazard function varies.

Each distribution has different characteristics and is suited to different types of data and analytical needs.

**[Exponential Distribution]**

Q\_1: What is the exponential distribution?

A\_1: The exponential distribution is a continuous probability distribution widely used in reliability engineering, queuing theory, survival analysis, and other fields. It is particularly useful for modeling the time intervals between events in a Poisson process, where events occur independently at a constant average rate. The distribution describes the time until the next event occurs.

Q\_2: What is the Poisson distribution?

A\_2: The Poisson distribution describes the number of events occurring within a fixed unit of time or space, assuming these events happen with a constant average rate and independently of one another. In contrast, the exponential distribution models the time between successive events. Specifically, given a Poisson process with an average rate of events denoted by $\lambda$, the number of events occurring in a fixed interval follows a Poisson distribution, while the time until the first event or between consecutive events follows an exponential distribution.

Q\_3: What is the expected value of the exponential distribution?

A\_3: The expected value (or mean) of the exponential distribution is given by: E\left(X\right) = 1/\lambda$

Q\_4: What is the variance of the exponential distribution?

A\_4: The variance of the exponential distribution is given by: $Var\left(X\right)= 1/\lambda^{2}$

Q\_5: What is the standard deviation of the exponential distribution?

A\_5: Therefore, the standard deviation is equal to the expected value.

Q\_6: What is the median of the exponential distribution?

A\_6: The median of the exponential distribution is given by: $0.693/\lambda$. This value is less than the expected value. The median being smaller than the mean indicates that the probability density function has a long right tail, meaning the distribution is asymmetric with a positive skew.

Q\_7: What is the skewness of the exponential distribution?

A\_7: Skewness measures the degree of asymmetry in a data distribution.

Negative skew: Indicates a longer left tail, with mass concentrated on the right side.

Positive skew: Indicates a longer right tail, with mass concentrated on the left side.

The exponential distribution has a positive skew, meaning it has a long right tail. This indicates that the distribution is asymmetric with values tending to be more spread out on the higher end compared to the mean.

Q\_8: What is the probability density function (PDF) of the exponential distribution?

A\_8: Exponential Distribution’s Probability Density Function $f(x, \lambda) = \lambda e^{-\lambda x} \left(x \geq 0\right) or 0 \left(x<0\right)$. Where x is the time until the event occurs, and e is a natural constant of about 2.7183. As x increases, the probability gradually decreases. Using the PDF, the probability of an event occurring within a specific time interval can be calculated. The probability of occurrence is obtained by integrating the PDF over the time interval desired to obtain the probability of occurrence.

Q\_9: What is the CDF of the exponential distribution?

A\_9: Since $\lambda$ is the rate parameter, $1/\lambda$ represents the average time between events. The probability that an event occurs within x hours (the cumulative distribution function, CDF) is given by: $P\left(X\leqx\right)=F\left(x\right)=\begin{cases}1-e^{-\lambda x} & x \geq0,\\0 & x < 0.\end{cases}$.

If the probability of an event occurring after x hours is desired, the cumulative distribution function (CDF) is given by:

$P\left(X\geq x\right)=1-F\left(x\right)=\begin{cases}e^{-\lambda x} & x \geq0,\\0 & x < 0.\end{cases}$

Q\_10: What is the memoryless property of the exponential distribution?

A\_10: The exponential distribution is memoryless. This means that the probability of an event occurring in a future time interval is independent of how much time has already passed. This property can be verified using conditional probability. Specifically, for positive real numbers t and s:

$P\left(X>t+s\mid X>t\right)=P\left(X>s\right)$

This equation shows that the probability of the event occurring after t+st + st+s hours, given that it has already occurred after t hours, is the same as the probability of the event occurring after s hours, regardless of t.

<Proof of the memoryless property of the exponential distribution>

PDF = $f\left(x\right)=\lambda e^{-\lambda x} for x\geq 0$, CDF = $F\left(x\right)=1-e^{-\lambda x}$ for $x\geq 0$. Calculating the conditional probability, we get: $P\left(X>t+s\mid X>t\right)=P\left(X>s\right)=\frac{P\left(X>t+s\cap X>t\right)}{P\left(X>t\right)}=\frac{P\left(X>t+s\right)}{P\left(X>t\right)}$ , For the exponential distribution, $P\left(X>t+s\right)=1-F\left(s+t\right)=e^{-\lambda\left(s+t\right)}$ , $P\left(X>s\right)=e^{-\lambda s}$. Therefore, $P\left(X>t+s\mid X>t\right)=\frac{e^{-\lambda\left(s+t\right)}}{e^{-\lambda s}}=e^{-\lambda}$. This is the same as $P\left(X>t\right)=e^{-\lambda t}$, so it satisfies inorganic repression.

Q\_11: What does "time intervals are independently distributed in the memoryless property of the exponential distribution" mean?

A\_11: The memoryless property means that the time you have already waited does not affect the time you will need to wait in the future. For example, if you have waited 10 minutes without receiving a call, the time you need to wait for the next call is independent of the 10 minutes you’ve already waited. It is as if you are starting from the beginning again, with the waiting time being determined randomly and independently of the past.

Q\_12: What is a Poisson process?

A\_12: A Poisson process models a sequence of events that occur independently and randomly over continuous time or space. To be considered a Poisson process, several conditions must be met:

1. Independence of Events: The number of events occurring in non-overlapping time intervals is independent of each other.

2. Constant Rate (Stationarity): The process has a constant average rate $\lambda$ at which events occur. This means that the average number of events in a given time interval is proportional to the length of the interval. For example, if you receive an average of 5 calls per hour, you would expect to receive 10 calls over two hours.

3. No Simultaneous Events: In a Poisson process, the probability of two or more events occurring at exactly the same time is zero.

Understanding the Poisson process helps in analyzing and predicting the occurrence of events in various real-life situations.

Q\_13: Can we use symbols other than $\lambda$ to represent the average rate of events in the exponential distribution?

A\_13: Yes, it is possible. $\lambda$ is a symbol representing the average rate of events, but you can use other symbols as well. The important thing is not the symbol itself, but its meaning. For example, you could use $\lambda$ or another symbol to represent the average rate of events, as long as it conveys that it represents the rate at which events occur.

Q\_14: What is conditional probability?

A\_14: Conditional probability is the probability of an event occurring given that another event has already occurred. For example, calculating the probability of receiving a call within the next 5 minutes given that no call has been received in the past 5 minutes is an example of conditional probability.

**[weibull distribution]**

Q\_1: What are the main differences between the Weibull distribution and the normal distribution or exponential distribution?

A\_1: The Weibull distribution is a flexible continuous probability distribution commonly used in life data analysis. It differs from the normal distribution and exponential distribution in that it can model scenarios where the failure rate either increases, decreases, or remains constant over time. The Weibull distribution has the flexibility to mimic other statistical distributions, such as the normal or exponential distributions, and is widely used in reliability analysis, failure analysis, and survival data analysis.

Q\_2: How do the scale parameter $\lambda$ and shape parameter k affect the behavior of the Weibull distribution?

A\_2: The scale parameter $\lambda$ determines the size and location of the distribution. As $\lambda$ increases, the distribution shifts to the right, indicating a higher likelihood of events with longer lifetimes. Conversely, a smaller $\lambda$ shifts the distribution to the left, indicating a higher likelihood of events with shorter lifetimes. The shape parameter k determines the form of the distribution. If k>1, the distribution is right-skewed, indicating an increasing failure rate over time. If k=1, the Weibull distribution becomes an exponential distribution, where the failure rate remains constant. If k<1, the distribution is left-skewed, indicating a decreasing failure rate over time.

Q\_3: What does each value of the shape parameter k>1, k=1, k<1 in the Weibull distribution mean?

A\_3: When k>1, the distribution is right-skewed, and the probability of failure increases over time. This is suitable for modeling situations where products initially function well but fail more frequently as time progresses.

When k=1, the Weibull distribution becomes an exponential distribution, where the probability of failure remains constant over time.

When k<1, the distribution is left-skewed, indicating that failures are more common early on, and the probability of failure decreases as time goes on.

Q\_4: In which areas is the Weibull distribution commonly used in industrial settings, and why is it preferred?

A\_4: The Weibull distribution is widely used in industrial settings for estimating component lifetimes, reliability analysis, and failure analysis. It is preferred because it can model situations where the failure probability increases, decreases, or remains constant over time. Additionally, its flexibility in adjusting scale and shape parameters allows it to mimic various statistical distributions, making it applicable to a wide range of scenarios.

Q\_5: How does the Weibull distribution model early failures in products, and what is the practical meaning of a left-skewed distribution?

A\_5: When k < 1 in the Weibull distribution, it is suitable for modeling products with frequent early failures. This situation often occurs with products that have initial defects or unstable quality, where the probability of failure is high initially but decreases over time. This distribution is useful in scenarios such as early failure analysis, warranty period setting, and evaluating the initial performance of medical equipment.

Q\_6: What are the differences between the 2-parameter Weibull distribution and the 3-parameter Weibull distribution, and when are each used?

A\_6: The 2-parameter Weibull distribution consists of only the scale parameter ($\lambda$) and the shape parameter (k), and it is used primarily for general lifetime data analysis. In contrast, the 3-parameter Weibull distribution includes an additional location parameter ($\gamma$), which allows for modeling scenarios where failures do not begin until after a certain amount of time has passed. For example, if a product operates without issues for a specific period before failures start, the 3-parameter Weibull distribution is used.

Q\_7: What is the formula for the probability density function (PDF) of the 2-parameter Weibull distribution, and how is it interpreted?

A\_7: The probability density function (PDF) of the 2-parameter Weibull distribution is expressed as: $f\left(x\beta,\eta\right)=\left(\frac{\beta}{\eta}\right)\times\left(\frac{x}{\eta}\right)^{\left(\beta-1\right)} \times\exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right]$

Here, $\beta$ is the shape parameter, and $\eta$ is the scale parameter. This formula represents the failure probability at a specific time x, and the shape of the distribution is determined by the values of $\eta$ and $\beta$.

Q\_8: How is the expected value (mean) of the Weibull distribution calculated, and what does it signify?

A\_8: The expected value (mean) of the 2-parameter Weibull distribution is calculated as follows: $E\left(X\right)=\eta\times\Gamma\left(1+\frac{1}{\beta}\right) $

Here, $\eta$ is the scale parameter, $\beta$. is the shape parameter, and $\Gamma$ is the gamma function. This value represents the mean lifetime and plays a crucial role in predicting the lifespan of products or systems.

Q\_9: Please explain about the wibble distribution.

A\_9: The Weibull distribution is a type of continuous probability distribution. It is named after Wallodi Weibull. In cases dealing with particle distributions, it is also referred to as the Rosin-Rammler distribution. The Weibull distribution is flexible and is frequently used in lifetime data analysis, as it can mimic other statistical distributions such as the normal distribution or the exponential distribution. It is a versatile probability distribution used in industry for estimating component lifetimes and is commonly used in reliability analysis, failure analysis, and survival data analysis. It can estimate cases where the failure probability increases, decreases, or remains constant over time. When the failure probability remains constant over time, it is similar to the exponential distribution.

The Weibull distribution has either 2 or 3 parameters.

Shape parameter: Determines the shape of the distribution.

Scale parameter: In probability theory and statistics, the scale parameter is a special type of mathematical parameter for a family of probability distributions. As the scale parameter increases, the distribution becomes more spread out.

Location parameter : A parameter that determines the location or shift of the distribution.

Q\_10: Tell me the characteristics of the scale parameter according to the size.

A\_10: The scale parameter $\lambda$ determines the position and size of the distribution.

If the scale parameter is large:

a) The overall size of the distribution increases.

b) Data shifts further to the right, and there is a higher likelihood of events with longer lifetimes occurring.

c) The tail of the distribution becomes wider, indicating an increased probability of extreme events.

d) A large scale parameter is often used to represent the lifetime data of reliable products or systems.

If the scale parameter is small:

a) The overall size of the distribution decreases.

b) Data shifts further to the left, and there is a higher likelihood of events with shorter lifetimes occurring.

c) The tail of the distribution becomes narrower, indicating a decreased probability of extreme events.

d) A small scale parameter is often used to represent the lifetime data of lower-quality or less reliable systems.

Q\_11: Explain the features for the shape parameter k.

A\_11: The characteristics of the shape parameter k are as follows.

For k > 1: Right-skewed distribution.

- When k is greater than 1, the distribution is skewed to the right.

- Products initially work well, but the likelihood of failure increases over time.

- This is useful for modeling situations where products endure for a long time before a sudden increase in failures.

For k = 1: Exponential distribution.

- When k = 1 , the Weibull distribution becomes an exponential distribution.

- Failures occur at a constant rate over time.

For k < 1: Left-skewed distribution.

- When k is less than 1, the distribution is skewed to the left.

- This represents products with a high failure rate early in their use, often seen with quality issues or initial defects.

- It is suitable for modeling situations where failures are more likely early on, and the likelihood of failure decreases over time.

**[Lognormal Distribution]**

Q\_1: Why is the expected value of a log-normal distribution not a simple arithmetic mean, but rather the mean of log-transformed values that is then exponentiated?

A\_1: In a log-normal distribution, the reason we don’t use a simple arithmetic mean for the expected value is that the distribution is not symmetric and has a long right tail. As a result, the expected value of X does not align with the mean in the log-transformed (normal) domain. The expected value is exponentiated to reflect the "geometric mean" of the original data, which is more appropriate for multiplicative processes. The geometric mean, unlike the arithmetic mean, provides a more representative central value when the data has a wide range, as is often the case with log-normal distributions.

Q\_2: What is log transformation?

A\_2: Log transformation is the process of replacing a variable X with its logarithmic value $\log\left(X\right)$. This transformation reduces the impact of large values and compresses the range of the data, often making the distribution more symmetric and linear, which can simplify analysis and improve the fit of certain statistical models.

Q\_3: What is the difference between a log-normal distribution and a normal distribution?

A\_3: A normal distribution allows for both positive and negative values, whereas a log-normal distribution cannot have negative values. In other words, values in a log-normal distribution must be greater than 0.

Q\_4: What is a log-normal distribution?

A\_4: A log-normal distribution is the distribution of a variable whose logarithm follows a normal distribution. Typically, this refers to the natural logarithm. In other words, if $\ln\left(X\right)$ follows a normal distribution, then X follows a log-normal distribution. It is also known as the Galton distribution.

Q\_5: What is the probability density function (PDF) of a log-normal distribution?

A\_5: The probability density function is as follows.

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

Conversely, if the variable X follows a normal distribution, then $\exp\left(X\right)$ follows a log-normal distribution.

Q\_6: Why is the log-normal distribution used?

A\_6: Variables following a log-normal distribution are always positive. Therefore, it serves as an alternative to the problems associated with normal distributions that can produce negative values. For instance, measurements like weight, density, height, energy, and length are always positive values.

Q\_7: What is a ‘random shock’?\*\*

A\_7: A ‘random shock’ refers to a phenomenon where independently occurring small changes multiply and affect the entire system.

Q\_8: What is "degradation phenomenon"?

A\_8: The "degradation phenomenon" refers to the process by which the performance, quality, or functionality of an object, system, or substance gradually deteriorates over time.

Q\_9: Define the log-normal distribution.  
A\_9: If a random variable Y follows a normal distribution Y∼N(μ,σ2)Y \sim $X\sim LN\left(\mu,\sigma^{2}\right)$, then the random variable $X = e^Y$ follows a log-normal distribution.

->In other words, if the logarithm of a random variable X follows a normal distribution with mean $,​ and variance $\sigma^{2}\_ {Y}$, then X follows a log-normal distribution with those parameters.

$X\sim LN\left(\mu,\sigma^{2}\right)$

$f(x \mid \mu, \sigma^2) = \frac{1}{x \sqrt{2\pi} \sigma} \exp\left( -\frac{(\log(x) - \mu)^2}{2\sigma^2} \right), \quad \text{if } 0 < x < \infty$

-Unlike a general normal distribution, the log-normal distribution can represent various types of distributions.

-This distribution is useful for modeling data that is skewed in the positive direction and cannot take negative values. Therefore, it is widely used as an empirical model for phenomena like failure data. Additionally, because it involves simply applying a logarithm to a normally distributed variable, it allows the use of various properties of normal distributions.

-It can be shown that the log-normal distribution arises from the cumulative effects of multiplicative shocks leading to failures.

Q\_10: What are the properties of the log-normal distribution?

A\_10: Here are the properties of the log-normal distribution:

-The log-normal distribution is defined only for x>0, as the exponential function cannot take negative values.

-The log-normal distribution is often used to model multiplicative processes, where independent random factors multiply to produce a result.

-If a variable X follows a log-normal distribution, then $\ln\left(x\right)$ follows a normal distribution. Conversely, if a variable Y follows a normal distribution, then follows a log-normal distribution.

Q\_11: What is the moment-generating function of the log-normal distribution?

A\_11: The moment-generating function of the log-normal distribution is derived from the normal distribution.

$M\_X(t) = E[e^{tX}] = E[e^{te^Y}] = \int\_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} \exp\left( te^Y - \frac{(y - \mu)^2}{2\sigma^2} \right) \, dy$

The expectation of the exponential function of a normal distribution is equivalent to computing the moment-generating function (mgf) for a random variable X that follows a log-normal distribution. This calculation cannot be simplified into a simple form.

Using the moment-generating function of the normal distribution, the moments of the log-normal distribution can be easily computed, which allows for the determination of its mean and variance.

Q\_12: How are the moments of the log-normal distribution calculated?

A\_12: The moments of the log-normal distribution are calculated as follows.

$M\_Y(t) = E[e^{tY}] = E[e^{t \log(x)}] = E[X^t] = e^{n\mu + \frac{n^2 \sigma^2}{2}}$

-> The moments of the log-normal distribution are obtained from the moment-generating function of the normal distribution by varying the value of 't'.

Q\_13: How can the mean of a log-normal distribution be calculated?

A\_13: The mean of a log-normal distribution can be calculated as follows:

$

$M\_Y(2) = E[e^{2Y}] = E[e^{2 \log(x)}] = E[X^2] = \exp\left(2\mu + 2\sigma^2\right)$

Q\_14: How can the variance of a log-normal distribution be calculated?

A\_14: The variance of a log-normal distribution can be calculated as follows:$\text{Var}(X) = M\_Y(2) - \left(M\_Y(1)\right)^2 = \exp(2\mu + \sigma^2) [\exp(\sigma^2) - 1]$

Q\_15: Why is the log-normal distribution often used as a model to simulate failures in real-world scenarios?

A\_15: When multiplicative small changes accumulate, applying a log transformation to use a normal distribution is equivalent to using a log-normal distribution. It is suitable when failure times follow a log-normal distribution in a logarithmic sense. Additionally, if degradation increases proportionally, failure times are modeled by a log-normal distribution.

Q\_16: What is 'multiplicative degradation'?

A\_16: Before explaining, consider the case where $X\_1, X\_2, \dots, X\_n$ are measurements of degradation at discrete time points. As X increases, it gets closer to failure, and when it reaches a certain level, failure occurs. (This threshold level varies from product to product.)

When the degradation phenomenon is assumed to be such that $x\_i = (1 + \epsilon\_i) x\_{i-1} $ ($\epsilon\_i$ (where $\epsilon\_i$​ are small, independent random 'shocks')

The degradation level of the subject increases by a factor of 1+$\epsilon\_i$ compared to the previous time point, and this is called "multiplicative degradation."

Q\_17: What is the relationship between "multiplicative degradation" and "log-normal distribution"?

A\_17: When the degradation level reaches a unique threshold for each product, failure occurs. Through this reasoning, we can understand that the lifespan of systems exhibiting "multiplicative degradation" tends to follow a log-normal distribution.

-The level of degradation at the n-th point in time from the initial point $x\_0$ can be calculated as follows:

$x\_n = \left[ \prod\_{i=1}^n (1 + \epsilon\_i) \right] x\_0$ ($x\_0$ is a constant, and $\epsilon\_i$ are random shocks.)

When taking the logarithm, it becomes as follows:

-> $\log(x\_n) = \sum\_{i=1}^n \log(1 + \epsilon\_i) + \log(x\_0) \approx \sum\_{i=1}^n \epsilon\_i + \log(x\_0)$

->By the Central Limit Theorem, $\log(x\_n)$ approximately follows a normal distribution. This implies that $x\_n$ approximately follows a lognormal distribution.

Therefore, since a failure occurs when the degradation level reaches a unique threshold for each product, we can deduce that the lifespan of systems exhibiting "multiplicative degradation" is well modeled by a lognormal distribution.

**[Method of distribution estimation Q&A]**

Q\_1: What is distribution estimation?

A\_1: Distribution estimation refers to the process of estimating which probability distribution a given dataset follows. Through this, the shape and characteristics of the data's distribution can be understood, and parameters such as mean, variance, and standard deviation can be estimated.

Q\_2: What is the K-S test?

A\_2: The K-S (Kolmogorov-Smirnov) test is a non-parametric statistical test that compares two cumulative distribution functions. It evaluates whether two datasets follow the same distribution or whether a given dataset follows a specific distribution by measuring the difference between the distributions.

Q\_3: What are the advantages of the K-S test?

A\_3: The advantages of the K-S test are:

1. It is non-parametric, making it flexible and applicable to various types of data.

2. It is relatively simple to compute and the results are easy to interpret.

3. It is based on the cumulative distribution function (CDF), making it effective at detecting overall differences in distributions.

4. It can provide reliable results even with small sample sizes.

5. It has strong test power compared to other non-parametric methods, making it sensitive to differences in data distributions.

Q\_4: What are the disadvantages of the K-S test?

A\_4: The disadvantages of the K-S test are:

1. It is not suitable for discrete data.

2. With large sample sizes, even small differences can appear significant, increasing the likelihood of rejecting the null hypothesis.

3. It is sensitive to extreme values, which can reduce the reliability of the results.

4. Accurately calculating the p-value for the K-S test can sometimes be complex, especially when dealing with large sample sizes.

Q\_5: How can the results of the K-S test be interpreted?

A\_5: In the K-S test, if the p-value is greater than 0.05, we do not reject the null hypothesis, meaning the data is likely to follow the assumed distribution. On the other hand, if the p-value is less than or equal to 0.05, we reject the null hypothesis, indicating the data is unlikely to follow the assumed distribution.

Q\_6: What is sensitivity?

A\_6: In statistics, sensitivity refers to a measure of how well a test detects true positives. A high sensitivity means that most, if not all, actual positive cases are correctly identified by the test.

Q\_7: What is likelihood?

A\_7: Likelihood refers to the probability that a model and its estimates fit the data well. It quantifies how well the model explains the observed data.

Q\_8: What is Maximum Likelihood Estimation (MLE)?

A\_8: Maximum Likelihood Estimation (MLE) is a method for estimating parameters of a probability density function $P(x \mid \theta)$ based on observed sample data. It involves finding the parameter values $\theta$ that maximize the likelihood function, which measures how well the observed data is explained by the model with those parameters.

Q\_9: Describe the process of Maximum Likelihood Estimation (MLE).

A\_9: The process of Maximum Likelihood Estimation (MLE) is as follows:

1. Define a probability model for the observed data and set up the likelihood function, which represents the probability of the observed data given the model parameters.

2. Compute the log-likelihood function by taking the logarithm of the likelihood function. The logarithm is used to simplify calculations and improve numerical stability.

3. Estimate the parameters by maximizing the log-likelihood function. This involves using differentiation or optimization algorithms to find the parameter values that yield the highest log-likelihood.

4. The parameter values that maximize the log-likelihood function are the MLE estimates.

Q\_10: What is the likelihood function?

A\_10: The likelihood function is a function that represents the probability of the observed data (x) given a specific parameter value ($\theta$). In other words, it quantifies the likelihood of the data occurring for a given parameter value. While the probability density function (PDF) calculates the probability of observing a given data point, the likelihood function calculates the probability of the observed data given a set of parameter values. The likelihood function is often denoted as

$L(\theta \mid x)$

Q\_11: What conditions are needed to have a good estimator?

A\_11: To have a good estimator, it should satisfy the following conditions: unbiasedness, efficiency, consistency, and sufficiency.

Q\_12: What does the property of unbiasedness in an estimator mean?

A\_12: Among the properties of an estimator, unbiasedness is as follows:

- Unbiasedness: An estimator is said to be unbiased if the expected value of the estimator equals the true parameter $\theta$.

- It measures whether, if an infinite number of samples were taken, the average of the estimates would converge to the true value.

- Although the estimates obtained from probabilistic samples may vary and have errors in individual cases, an unbiased estimator ensures that, on average, these estimates will be correct and match the true parameter value.

Q\_13: What does the property of efficiency in an estimator mean?

A\_13: The property of efficiency in an estimator means the following:

- Efficiency: It refers to having a small mean squared error (MSE) for the parameter estimator. A smaller MSE indicates that the estimator's error is generally small on average, making it a good estimator.

- If the estimator is unbiased, its MSE consists only of the variance inherent to the estimator.

- When comparing unbiased estimators, only the variance of the estimators is compared. However, when comparing biased estimators or between biased and unbiased estimators, the mean squared error (MSE) is used for comparison.

Q\_14: What does consistency of an estimator mean?

A\_14: The property of consistency of an estimator is as follows:

- Consistency refers to the property where, as the sample size increases, the point estimator yields values closer to the true parameter value.

- Given an estimator $\hat{\theta}\_n$ for a sample size n, if for any $\epsilon > 0$, $\Pr \left( \left| \hat{\theta}\_n - \theta \right| < \epsilon \right)^{n\rightarrow\infty}$, then $\hat{\theta}\_n$ is called a 'consistent estimator' for the parameter $\theta$.

Q\_15: What does sufficiency of an estimator mean?

A\_15: The property of sufficiency of an estimator is as follows:

- Sufficiency refers to a statistic's ability to capture all the information about a parameter contained in the sample data.

Q\_16: What are the differences between non-parametric tests and parametric tests?

A\_16: The differences between non-parametric and parametric tests are:

Parametric Tests

- Assumes a specific distribution for the population.

- Makes inferences about the population based on statistics obtained from a sample.

- Examples include t-tests, paired t-tests, and ANOVA.

Non-parametric Tests

- Does not assume a specific distribution for the population.

- Used when the conditions for parametric tests are not met or when the sample size is small.

- Uses median instead of mean for the central tendency of the sample.

Q\_17: What is a one-sample K-S test, and how is it performed?

A\_17: The one-sample K-S test is used to determine whether a given dataset follows a specific cumulative distribution function (CDF). The procedure for performing a one-sample K-S test is as follows:

<One-sample K-S Test Procedure>

1. Sort the Data: Arrange the data from the smallest to the largest value.

2. Calculate CDF: Compute the empirical cumulative distribution function (ECDF) of the data and compare it with the CDF of the distribution being tested.

3. Compute Maximum Difference: Calculate the maximum difference between the two CDFs to obtain the K-S statistic D .

4. Calculate p-value: Compute the p-value corresponding to the D statistic and decide whether to reject the null hypothesis.

Q\_18: What is a two-sample K-S test, and how is it performed?

A\_18: The two-sample K-S test is used to determine whether two independent datasets follow the same distribution. The procedure for performing a two-sample K-S test is as follows:

<Two-sample K-S Test Procedure>

1. Sort the Data: Sort each of the two datasets individually.

2. Calculate CDF: Compute the empirical cumulative distribution function (ECDF) for each dataset.

3. Compute Maximum Difference: Calculate the maximum difference between the two ECDFs to obtain the K-S statistic D .

4. Calculate p-value: Compute the p-value corresponding to the D statistic and determine whether the two datasets follow the same distribution.