1.010 - Brief Notes # 8
Selected Distribution Models

• The Normal (Gaussian) Distribution:

Let $X_1, \ldots, X_n$ be independent random variables with common distribution $F_X(x)$. The so called central limit theorem establishes that, under mild conditions on $F_X$, the sum $Y = X_1 + \ldots + X_n$ approaches as $n \to \infty$, a limiting distributional form that does not depend on $F_X$. Such a limiting distribution is called the Normal or Gaussian distribution. It has the probability density function:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(y-m)^2/\sigma^2}$$

where

$m = \text{mean value of } Y$

$\sigma = \text{standard deviation of } Y$

Notice: $m = nm_X$, $\sigma^2 = n\sigma_X^2$, where $m_X$ and $\sigma_X^2$ are the mean value and variance of $X$.

• Properties of the Normal (Gaussian) Distribution:

1. For most distributions $F_X$, convergence to the normal distribution is obtained already for $n$ as small as 10.

2. Under mild conditions, the distribution of $\sum_i X_i$ approaches the normal distribution also for dependent and differently distributed $X_i$.

3. If $X_1, \ldots, X_n$ are independent normal variables, then any linear function $Y = a_0 + \sum_i a_i X_i$ is also normally distributed.
• **The Lognormal Distribution:**

Let \( Y = W_1 W_2 \cdots W_n \), where the \( W_i \) are iid, positive random variables. Consider:

\[
X = \ln Y = \sum_{i=1}^{n} \ln W_i
\]

For \( n \) large, \( X \sim N(m_{\ln Y}, \sigma_{\ln Y}^2) \)

\( Y = e^X \) has a lognormal distribution with PDF:

\[
f_Y(y) = \frac{dx}{dy} f_X(x(y)) = \frac{1}{y} \frac{1}{\sqrt{2\pi} \sigma_{\ln Y}} e^{-\frac{1}{2} \left( \ln y - m_{\ln Y} \right)^2 / \sigma_{\ln Y}^2}, \quad y \geq 0
\]
If $X \sim N(m_X, \sigma_X^2)$, then $Y = e^X \sim LN(m_Y, \sigma_Y^2)$ with mean value and variance given by:

$$
\begin{align*}
    m_Y &= e^{m_X + \frac{1}{2}\sigma_X^2} \\
    \sigma_Y^2 &= e^{2m_X + \sigma_X^2} \left(e^{\sigma_X^2} - 1\right)
\end{align*}
$$

Conversely, $m_X$ and $\sigma_X^2$ are found from $m_Y$ and $\sigma_Y^2$ as follows:

$$
\begin{align*}
    m_X &= 2\ln(m_Y) - \frac{1}{2}\ln(\sigma_Y^2 + m_Y^2) \\
    \sigma_X^2 &= -2\ln(m_Y) + \ln(\sigma_Y^2 + m_Y^2)
\end{align*}
$$

Property: products and ratios of independent lognormal variables are also lognormally distributed.
• **The Beta Distribution:**

The Beta distribution is commonly used to describe random variables with values in a finite interval. The interval may be normalized to be $[0, 1]$. The Beta density can take on a wide variety of shapes. It has the form:

$$f_Y(y) \propto y^a(1 - y)^b$$

where $a$ and $b$ are parameters. For $a = b = 0$, the Beta distribution becomes the uniform distribution.

![Standard Beta PDF](image)

• **Multivariate Normal Distribution:**

Consider $Y = \sum_{i=1}^{n} X_i$, where the $X_i$ are iid random vectors.

As $n$ becomes large, the joint probability density of $Y$ approaches a form of the type:

$$f_Y(y) = \frac{(\text{det } \Sigma)^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(y - \mu)^T \Sigma^{-1} (y - \mu)}$$

where $\mu$ and $\Sigma$ are the mean vector and covariance matrix of $Y$. 
Properties

1. Contours of $f_Y$ are ellipsoids centered at $m$.
2. If the components of $Y$ are uncorrelated, then they are independent.
3. The vector $Z = a + BY$, where $a$ is a given vector and $B$ is a given matrix, has jointly normal distribution $N(a + Bm, B\Sigma B^T)$.

Let \[
\begin{bmatrix}
Y_1 \\
Y_2
\end{bmatrix}
\]
be a partition of $Y$, with associated partitioned mean vector \[
\begin{bmatrix}
m_1 \\
m_2
\end{bmatrix}
\]
and covariance matrix \[
\begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix}
\]. Then:

4. $Y_i$ has jointly normal distribution: $N(m_i, \Sigma_{ii})$.
5. $(Y_1 \mid Y_2 = y_2)$ has normal distribution $N(m_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - m_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

The bivariate normal distribution.
Relationships between Mean and Variance of Normal and Lognormal Distributions

If \( X \sim N(m_X, \sigma_X^2) \), then \( Y = e^X \sim LN(m_Y, \sigma_Y^2) \) with mean value and variance given by:

\[
\begin{align*}
  m_Y &= e^{m_X + \frac{1}{2}\sigma_X^2} \\
  \sigma_Y^2 &= e^{2m_X + \sigma_X^2} \left( e^{\sigma_X^2} - 1 \right)
\end{align*}
\]

Conversely, \( m_X \) and \( \sigma_X^2 \) are found from \( m_Y \) and \( \sigma_Y^2 \) as follows:

\[
\begin{align*}
  m_X &= 2 \ln(m_Y) - \frac{1}{2} \ln(\sigma_Y^2 + m_Y^2) \\
  \sigma_X^2 &= -2 \ln(m_Y) + \ln(\sigma_Y^2 + m_Y^2)
\end{align*}
\]