Recall the definition of a probability space: a triple \((\Omega, \mathcal{B}, P)\) where \(\Omega\) is a set, \(\mathcal{B}\) is a set of subsets of \(\Omega\) (a \(\sigma\)-algebra) such that

- \(\emptyset \in \mathcal{B}\)
- if \(B \in \mathcal{B}\) then \(\Omega \setminus B \in \mathcal{B}\) (so \(\Omega = \Omega \setminus \emptyset \in \mathcal{B}\))
- if \(B_1, B_2, \ldots \in \mathcal{B}\) then \(\bigcup_{i=1}^{\infty} B_i \in \mathcal{B}\)

and \(P : \mathcal{B} \to [0,1]\) (the probability) is a function such that

- \(P(\Omega) = 1\)
- if \(B_1, B_2, \ldots\) are disjoint elements of \(B\), then

\[
P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i)
\]

(countable additivity).
Random variables

A (real-valued) random variable $X$ is a function $X : \Omega \to \mathbb{R}$ such that for all $b \in \mathbb{R}$, the set

\[ \{ \omega \in \Omega : X(\omega) < b \} \in \mathcal{B} \]

The cumulative distribution function $F_X$ of $X$ is the function $F_X : \mathbb{R} \to \mathbb{R}$ given by

\[ F_X(x) = P(\{ \omega \in \Omega : X(\omega) \leq x \}) \]

Often in practice, $F_X$ is differentiable, and then

\[ f_X = F_X' \]

is called the probability density function, and then

\[ F_X(x) = \int_{-\infty}^{x} f_X(t) \, dt \]

The expected value of $X$ is

\[ E[X] = \int_{-\infty}^{\infty} t \, f_X(t) \, dt = \int_{-\infty}^{\infty} t \, dF_X(t) \]

and the variance of $X$ is

\[ \sigma^2 = \sigma^2[X] = \text{Var}[X] = \int_{-\infty}^{\infty} (t - E[X])^2 \, dF_X(t) \]
The normal distribution with parameters 
\[ \mu \] (mean, or expectation) and 
\[ \sigma \] (standard deviation) has the cumulative distribution function
\[
N_{\mu, \sigma}(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2} \left( \frac{t-\mu}{\sigma} \right)^2} dt
\]
and the probability density function
\[
\phi(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}.
\]

Thus our \( N(x) \) is the same as \( N_{\mu, \sigma}(x) \).

If a random variable has the distribution given by \( N_{\mu, \sigma}(x) \) (i.e., if \( F_X(x) = N_{\mu, \sigma}(x) \) for all \( x \in \mathbb{R} \)), we write
\[ X \sim N(\mu, \sigma^2). \]

In what follows, we often consider a random variable \( Z \) such that
\[ Z \sim N(0,1). \]
If \( X \sim \mathcal{N}(\mu, \sigma^2) \), and if we define
\[
Z = \frac{X - \mu}{\sigma},
\]
then
\[
Z \sim \mathcal{N}(0, 1).
\]
Thus for \( b \in \mathbb{R} \),
\[
P(\{X \leq b\}) = P(\{ \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma} \})
\]
\[
= P(\{Z \leq \frac{b - \mu}{\sigma} \}) = N\left(\frac{b - \mu}{\sigma}\right).
\]
Also,
\[
P(\{X > b\}) = 1 - P(\{X \leq b\})
\]
\[
= 1 - N\left(\frac{b - \mu}{\sigma}\right) = N\left(-\frac{b - \mu}{\sigma}\right) = N\left(\frac{\mu - b}{\sigma}\right)
\]
Since \( N(x) + N(-x) = 1 \) for all \( x \in \mathbb{R} \).
If \( X_1, X_2 \) are random variables with 
\[
E(X_i) = \mu_i, \quad \text{Var}(X_i) = \sigma_i^2,
\]
then the correlation between \( X_1 \) and \( X_2 \) is given by
\[
P_{12} = \frac{1}{\sigma_1 \sigma_2} E[(X_1 - \mu_1)(X_2 - \mu_2)].
\]
The covariance of \( X_1, X_2 \) is \( \text{Cov}[X_1, X_2] = \sigma_{12} = P_{12} \sigma_1 \sigma_2 \).

If \( X_1, \ldots, X_n \) are random variables,
\[
E\left[ \sum_{i=1}^{n} a_i X_i \right] = \sum_{i=1}^{n} a_i E[X_i] = \sum_{i=1}^{n} a_i \mu_i,
\]
\[
\text{Var}\left[ \sum_{i=1}^{n} a_i X_i \right] = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \sigma_{ij}.
\]

The normal distribution has the special property that if \( X_i \sim N(\mu_i, \sigma_i^2) \), then
\[
\sum_{i=1}^{n} a_i X_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \sigma_{ij}),
\]
provided that \( X_1, \ldots, X_n \) are also jointly normally distributed.
The Central Limit Theorem

Let $X_1, X_2, \ldots$ be random variables with the same expected value $\mu$ and the same variance $\sigma^2$. Suppose that the $X_i$ are independent (if $i \neq j$) and $\sigma, \mu \in \mathbb{R}$ then

$$P(\{X_i \leq a \text{ and } X_j \leq b\}) = P(\{X_i \leq a\})P(\{X_j \leq b\})$$

and identically distributed ($F_i = F_j$).

Then, as $n \to \infty$, we have

$$\frac{S_n - \mu}{\sigma \sqrt{n}} \to N(0, 1)$$

In this sense,

$$\frac{S_n - \mu}{\sigma \sqrt{n}} \text{ tends to the standard normal distribution } N(0, 1).$$

There are also stronger forms of the central limit theorem where the $\mu$ and $\sigma$ need not be the same for all $X_i$ (for example).
Exponentials of certain random variables.

Let $X$, $Y$ be random variables such that

$$Y = e^X,$$

so $Y > 0$ and

$$X = \ln Y.$$

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$E[Y] = E[e^X] = e^{\mu + \frac{1}{2} \sigma^2}$$

and

$$\text{Var}[Y] = \text{Var}[e^X] = e^{2\mu + \sigma^2} \left( e^\sigma^2 - 1 \right).$$

Note that this is based on the assumption that $X \sim \mathcal{N}(\mu, \sigma^2)$. 