Remark. The definition of mutual independence of random variables is same if the $X_i$ are random vectors if one makes the appropriate changes to the dimensions of the subvectors $A_i$. Also, the theorem about $g_1(X_1) \rightarrow g_n(X_n)$ also holds, with the same proof, if one changes the function $g_i$ to be from $R^{d_i} \rightarrow R^{d'_i}$, where $d_i$ is the dimension of $X_i$ and $d'_i \leq d_i, i = 1, \ldots, n$.

Eq. Say $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ and $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. Then $X$ and $Y$ are independent if $P(X \in A_1, Y \in A_2) = P(X \in A_1)P(Y \in A_2)$, for all $A_1, A_2$ with $A_1 \subset R^3$ and $A_2 \subset R^2$. Say $X$ and $Y$ are independent. Let $g_1(X) = X_1 + X_2 + X_3$ and $g_2(Y) = \begin{bmatrix} Y_1 \\ Y_2 - Y_1 \end{bmatrix}$. Then $g_1(X)$ and $g_2(Y)$ are independent.

Theorem. Suppose $X_1, \ldots, X_n$ are $n$ random variables that are either jointly discrete with joint pmf $p_X(x_1, \ldots, x_n)$ or jointly continuous with joint pdf $f_X(x_1, \ldots, x_n)$. Also, let $p_{X_i}(x_i)$ be the marginal pmf of $X_i$ if $X_1, \ldots, X_n$ are jointly discrete, and let $f_{X_i}(x_i)$ be the marginal pdf of $X_i$ if $X_1, \ldots, X_n$ are jointly continuous. Then $X_1, \ldots, X_n$ are mutually independent if and only if $p_X(x_1, \ldots, x_n) = p_{X_1}(x_1) \times \cdots \times p_{X_n}(x_n)$ (discrete case) or $f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)$ (continuous case).

For all $(x_1, \ldots, x_n)^T \in R^n$.

Proof. Suppose that the factoring in (2) holds. Let $A_1, \ldots, A_n \subset R$ be arbitrary. Let $S_x$ denote the support of the distribution of the random vector $(X_1, \ldots, X_n)^T$. Then
\[
P(X_1 \in A_1, \ldots, X_n \in A_n) = P((X_1, \ldots, X_n)^T \in A_1 \times \ldots \times A_n)
\]
\[
= \sum_{x_n \in A_n \cap \mathbb{R}^n} \sum_{x_i \in A_i \cap \mathbb{R}^n} p(x_1, \ldots, x_n) \quad \text{(discrete case)}
\]
\[
= \int_{A_n} \ldots \int_{A_1} f(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \quad \text{(continuous case)}
\]
\[
= \sum_{x_n \in A_n \cap \mathbb{R}^n} \sum_{x_i \in A_i \cap \mathbb{R}^n} p_{X_i}(x_i) \times \ldots \times p_{X_n}(x_n) \quad \text{(discrete case)}
\]
\[
= \int_{A_n} \ldots \int_{A_1} f_{X_i}(x_i) \times \ldots \times f_{X_n}(x_n) \, dx_1 \ldots dx_n \quad \text{(continuous case)}
\]
\[
= \prod_{i=1}^{n} p_{X_i}(x_i) \times \prod_{i=1}^{n-1} p_{X_{i+1}}(x_{i+1}) 
\]
\[
= P(X_n \in A_n) \times P(X_{n-1} \in A_{n-1}) \times \ldots \times P(X_1 \in A_1)
\]

Since \(A_1, \ldots, A_n\) were arbitrary, \(X_1, \ldots, X_n\) are mutually independent.

Next, suppose that \(X_1, \ldots, X_n\) are mutually independent.

**Discrete Case**

Let \((x_1, \ldots, x_n)^T \in \mathbb{R}^n\) be arbitrary. By mutual independence,
\[
P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \times \ldots \times P(X_n = x_n)
\]
\[
p_{X_i}(x_i) \times \ldots \times p_{X_n}(x_n)
\]

**Continuous Case**

Let \((x_1, \ldots, x_n)^T \in \mathbb{R}^n\) be arbitrary. By mutual independence,
\[
P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \times \ldots \times P(X_n \leq x_n)
\]
This implies that \(F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \times \ldots \times F_{X_n}(x_n)\), where
\(F_X\) is the joint df of \((X_1, \ldots, X_n)^T\) and \(F_{X_i}\) is the marginal df of \(X_i\), \(i = 1, \ldots, n\).
Then
\[ f_X(x_1, \ldots, x_n) = \frac{d^n}{dx_1 \ldots dx_n} F_X(x_1, \ldots, x_n) \]
\[ = \frac{d^n}{dx_1 \ldots dx_n} \left( F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \right) \]
\[ = \left( \frac{d}{dx_1} F_{X_1}(x_1) \right) \times \cdots \times \left( \frac{d}{dx_n} F_{X_n}(x_n) \right) \]
\[ = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \]

**Theorem** Let \( F_X(x_1, \ldots, x_n) \) be the joint df of \((X_1, \ldots, X_n)^T\) and let \( F_{X_i}(x_i) \) be the marginal df of \( X_i \), \( i = 1, \ldots, n \). Then \( X_1, \ldots, X_n \) are mutually independent if and only if
\[ F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \]

**Proof** We already showed in the last part of the proof of the last theorem that if \( X_1, \ldots, X_n \) are mutually independent, then
\[ F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \]
Now, suppose that \( F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \) holds for all \((x_1, \ldots, x_n)^T \in \mathbb{R}^n\). Then
\[ P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \times \cdots \times P(X_n \leq x_n) \text{ for all } (x_1, \ldots, x_n) \in \mathbb{R}^n. \]
So the condition in the definition of mutual independence holds for subsets \( A_i \) of the form \( A_i = (-\infty, x_i] \), \( i = 1, \ldots, n \). What we need to show is that if
\[ P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \times \cdots \times P(X_n \leq x_n) \text{ for all } (x_1, \ldots, x_n) \in \mathbb{R}^n, \]
then
\[ P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \times \cdots \times P(X_n \in A_n). \]
This is outside the scope of this course.