Linear Algebra & Random Processes

Multiple Random Variables

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Topics Covered & References

Topics

References

Multiple Random Variables

- In many experiments, it is not uncommon for us to be interested in more than one variable related to the experiment. For example, if we were collecting data on obesity in a population, we would measure the age, weight, height, dietary habits etc. of each individual participant in the study.
- Thus, often our sample space of interest S is a Cartesian product of "smaller" sample spaces $S_1, S_2, \dots S_n$: $S = S_1 \times S_2 \times \cdots \times S_n$. Our probability model will now have n random variables, one for each S_i, and our pmf/pdfs will be multi-variate functions. We define a n-dimensional random vector as an element of $\mathbb{R}^n$ which can be compactly written as the following,

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^T \in \mathbb{R}^n$$

Multiple Random Variables - Discrete Case

- Consider the bivariate case, $\mathbf{z} == \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$. We can define the joint probability mass function $\mathbf{x}$ as the following,

$$f_{\mathbf{x},\mathbf{y}}\left(x, y\right) = P\left(\mathbf{x} = x, \mathbf{y} = y\right) = P\left(\mathbf{x} = x \cap \mathbf{y} = y\right)$$

- Probabilities of events defined on $\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}$ can be determined through the following,

$$P\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in A \subset \mathbb{R}^2\right) = \sum_{[x,y]^T \in A} f_{\mathbf{x},\mathbf{y}}\left(x, y\right)$$

- Joint probability distribution function of multivariate r.v. is defined as,

$$F_{\mathbf{x},\mathbf{y}}\left(x, y\right) = P\left(\mathbf{x} \leq x, \mathbf{x} \leq y\right) = \sum_{u \leq x} \sum_{v \leq y} f_{\mathbf{x},\mathbf{y}}\left(u, v\right)$$

Multiple Random Variables - Discrete Case

Marginal Probability Mass Functions – the pmf of the the individual r.v.s.

$$f_{\mathbf{x}_1}(x_1) = \sum_{x_2} f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2)$$

$$f_{\mathbf{x}_2}(x_2) = \sum_{x_1} f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2)$$

$f_{\mathbf{x}_1,\mathbf{x}_1}(x_1, x_2)$		$\mathbf{x}_1$						$f_{\mathbf{x}_2}(x_2)$
		1	2	3	4	5	6	
$\mathbf{x}_2$	1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
$f_{\mathbf{x}_1}(x_1)$		$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Multiple Random Variables - Discrete Case

Marginal Probability Mass Functions – the pmf of the the individual r.v.s.

$$f_{\mathbf{x}_1}(x_1) = \sum_{x_2} f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2)$$

$$f_{\mathbf{x}_2}(x_2) = \sum_{x_1} f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2)$$

$f_{\mathbf{x}_1,\mathbf{x}_1}(x_1, x_2)$		$\mathbf{x}_1$						$f_{\mathbf{x}_2}(x_2)$
		1	2	3	4	5	6	
$\mathbf{x}_2$	1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
	6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
$f_{\mathbf{x}_1}(x_1)$		$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Consider a pair of dice that are thrown, and let $\mathbf{x}_1$ and $\mathbf{x}_2$ be random variables representing the number that turns up on each of the two dice.s What will be the PMFs of the following bivariate r.vs $\mathbf{z} =$ (a) $[\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2]^T$; (c) $[\mathbf{x}_1 \cdot \mathbf{x}_2, \mathbf{x}_1 / \mathbf{x}_2]^T$.

Multiple Random Variables - Continuous Case

- The joint pdf for the bivariate case gives us the probability of the r.v.s assuming a value in a small interval around the point of interest.
$$P\left(x - \frac{\delta x}{2} < \mathbf{x} \leq x + \frac{\delta x}{2}, y - \frac{\delta y}{2} < \mathbf{y} \leq y + \frac{\delta y}{2}\right) = f_{\mathbf{x},\mathbf{y}}\left(x, y\right) \delta x \delta y$$
- Probabilities of events defined on $\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}^T$ can be determined through the following,
$$P\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in A \subset \mathbb{R}^2\right) = \iint_A f_{\mathbf{x},\mathbf{y}}\left(x, y\right) dx\, dy$$
- Joint probability distribution function of multivariate r.v. is defined as,
$$F_{\mathbf{x},\mathbf{y}}\left(x, y\right) = P\left(\mathbf{x} \leq x, \mathbf{x} \leq y\right) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{\mathbf{x},\mathbf{y}}\left(u, v\right) du\, dv$$

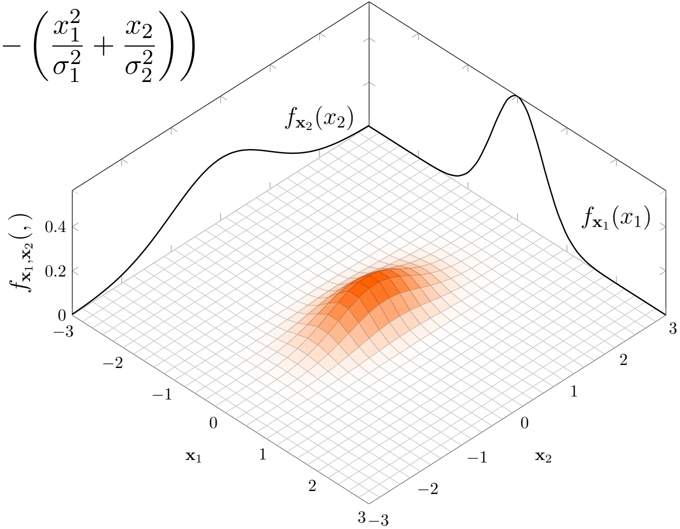
Multiple Random Variables - Continuous Case

$$f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\left(\frac{x_1^2}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}\right)\right)$$

$$f_{\mathbf{x}_1} = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x_1^2}{\sigma_1^2}\right)$$

$$f_{\mathbf{x}_2} = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{x_2^2}{\sigma_2^2}\right)$$

This is an example of a multivariate Guassian distribution. In this case the marginal distributions are also Gaussian.



Conditional Distributions and Independence

- Often, the knowledge of one r.v. in a set of r.v.s gives us some information about the other r.v.s.
- Consider a discrete bivariate r.v. $\mathbf{z} = [\mathbf{x},\ \mathbf{y}]^T$, with pmf $f_{\mathbf{x},\mathbf{y}}\left(x, y\right)$ and marginal pmfs $f_{\mathbf{x}}\left(x\right)$ and $f_{\mathbf{y}}\left(y\right)$. The *conditional pmf* of $\mathbf{y}$ given that $\mathbf{x} = x$ is given by,

$$f_{\mathbf{y}|\mathbf{x}}\left(y|x\right) = P\left(y|x\right) = \frac{P\left(\mathbf{x} = x, \mathbf{y} = y\right)}{P\left(\mathbf{x} = x\right)} = \frac{f_{\mathbf{x},\mathbf{y}}\left(x, y\right)}{f_{\mathbf{x}}\left(x\right)}$$

- We can similarly define the condition pdf of a continuous bivariate r.v. as,

$$f_{\mathbf{y}|\mathbf{x}}\left(y|x\right) = P\left(y|x\right) = \frac{P\left(\mathbf{x} = x, \mathbf{y} = y\right)}{P\left(\mathbf{x} = x\right)} =$$

Conditional Distributions and Independence

- Often, the knowledge of one r.v. in a set of r.v.s gives us some information about the other r.v.s.
- Consider a discrete bivariate r.v. $\mathbf{z} = [\mathbf{x}, \ \mathbf{y}]^T$, with pmf $f_{\mathbf{x},\mathbf{y}}(x, y)$ and marginal pmfs $f_{\mathbf{x}}(x)$ and $f_{\mathbf{y}}(y)$. The *conditional pmf* of $\mathbf{y}$ given that $\mathbf{x} = x$ is given by,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

- We can similarly define the condition pdf of a continuous bivariate r.v. as,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

Conditional Distributions and Independence

- Often, the knowledge of one r.v. in a set of r.v.s gives us some information about the other r.v.s.
- Consider a discrete bivariate r.v. $\mathbf{z} = [\mathbf{x},\ \mathbf{y}]^T$, with pmf $f_{\mathbf{x},\mathbf{y}}(x, y)$ and marginal pmfs $f_{\mathbf{x}}(x)$ and $f_{\mathbf{y}}(y)$. The *conditional pmf* of $\mathbf{y}$ given that $\mathbf{x} = x$ is given by,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

- We can similarly define the condition pdf of a continuous bivariate r.v. as,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

Consider a pair of dice that are thrown, and let $\mathbf{x}_1$ and $\mathbf{x}_2$ be r.v.s representing the number on each dice. Let $\mathbf{z} = [\mathbf{z}_1,\ \mathbf{z}_2]^T = [\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2]^T$. What is $f_{\mathbf{z}_2|\mathbf{z}_1}(z_2|z_1 = 4)$?

Conditional Distributions and Independence

- Often, the knowledge of one r.v. in a set of r.v.s gives us some information about the other r.v.s.
- Consider a discrete bivariate r.v. $\mathbf{z} = [\mathbf{x},\ \mathbf{y}]^T$, with pmf $f_{\mathbf{x},\mathbf{y}}(x, y)$ and marginal pmfs $f_{\mathbf{x}}(x)$ and $f_{\mathbf{y}}(y)$. The *conditional pmf* of $\mathbf{y}$ given that $\mathbf{x} = x$ is given by,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

- We can similarly define the condition pdf of a continuous bivariate r.v. as,

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = P(y|x) = \frac{P(\mathbf{x} = x, \mathbf{y} = y)}{P(\mathbf{x} = x)} = \frac{f_{\mathbf{x},\mathbf{y}}(x, y)}{f_{\mathbf{x}}(x)}$$

Consider a pair of dice that are thrown, and let $\mathbf{x}_1$ and $\mathbf{x}_2$ be r.v.s representing the number on each dice. Let $\mathbf{z} = [\mathbf{z}_1,\ \mathbf{z}_2]^T = [\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2]^T$. What is $f_{\mathbf{z}_2|\mathbf{z}_1}(z_2|z_1 = 4)$?

Consider a joint pdf $f_{\mathbf{x},\mathbf{y}}(x, y) = e^{-y}, 0 < x < y < \infty$. What is $f_{\mathbf{y}|\mathbf{x}}(y|x)$? $f_{\mathbf{x}|\mathbf{y}}(x|y)$?

Transformation of multiple random variables

- We can extend the idea of transformation of r.v.s to the case with multiple random variables.
- Consider the simple case of where two r.v.s $\mathbf{x}_1, \mathbf{x}_2$ are mapped to a single r.v. $\mathbf{y}$,

$$\mathbf{y} = g\left(\mathbf{x}_1, \mathbf{x}_2\right) \longrightarrow F_{\mathbf{y}}\left(y\right) = P\left(\mathbf{y} \leq y\right) = \iint_{x_1, x_2 \in D_y} f_{\mathbf{x}_1, \mathbf{x}_2}\left(x_1, x_2\right) dx_1\, dx_2$$

When, $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2$ and $f_{\mathbf{x}_1, \mathbf{x}_2}\left(x_1, x_2\right) = f_{\mathbf{x}_1}\left(x_1\right) f_{\mathbf{x}_2}\left(x_2\right)$,

$$f_{\mathbf{y}}\left(y\right) = \int_{x_1=-\infty}^{\infty} f_{\mathbf{x}_1}\left(x_1\right) f_{\mathbf{x}_2}\left(y - x_1\right) dx = \int_{x_2=-\infty}^{\infty} f_{\mathbf{x}_2}\left(y - x_2\right) f_{\mathbf{x}_2}\left(x_2\right) dx_2$$

For the discrete case, we have

$$f_{\mathbf{y}}\left(y\right) = \sum_{x_1=-\infty}^{\infty} f_{\mathbf{x}_1}\left(x_1\right) f_{\mathbf{x}_2}\left(y - x_1\right) = \sum_{x_2=-\infty}^{\infty} f_{\mathbf{x}_1}\left(y - x_2\right) f_{\mathbf{x}_2}\left(x_2\right)$$

Transformation of multiple random variables

- Lets now consider a general transformation that maps multiple r.v.s to multiple r.v.s, i.e. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$ to $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$, such that $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} g_1\left(\mathbf{x}_1, \mathbf{x}_2\right) \\ g_2\left(\mathbf{x}_1, \mathbf{x}_2\right) \end{bmatrix}$
$$F_{\mathbf{y}_1, \mathbf{y}_2}\left(y_1, y_2\right) = P\left(\mathbf{y}_1 \leq y_1, \mathbf{y}_2 \leq y_2\right) = \iint\limits_{x_1, x_2 \in D_{y_1, y_2}} f_{\mathbf{x}_1, \mathbf{x}_2}\left(x_1, x_2\right) dx_1\, dx_2$$
- The joint density function $f_{\mathbf{y}_1, \mathbf{y}_2}\left(y_1, y_2\right)$ is related to $f_{\mathbf{x}_1, \mathbf{x}_2}\left(x_1, x_2\right)$ through the Jacobian of the map from $\mathbf{x}$ to $\mathbf{y}$.
$$f_{\mathbf{y}_1, \mathbf{y}_2}\left(y_1, y_2\right) = \frac{1}{\left|\mathbf{J}\left(x_1, x_2\right)\right|} f_{\mathbf{x}_1, \mathbf{x}_2}\left(x_1, x_2\right)$$

where, $\mathbf{J}\left(x_1, x_2\right) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{bmatrix}$.

Expected values of multi-variate random variables

- Consider a n-dimensional r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^\top$, and a function $\mathbf{y} = g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)$. The expected value of this function is given by,

$$\mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) f_{\mathbf{x}_1, \ldots \mathbf{x}_n}\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) dx_1 \ldots dx_n$$

E.g. $g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) = \mathbf{x}_1 + \cdots + \mathbf{x}_n \implies \mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \sum_{i=1}^{n} \mathbb{E}\left(\mathbf{x}_i\right)$

Expected values of multi-variate random variables

- Consider a n-dimensional r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^\top$, and a function $\mathbf{y} = g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)$. The expected value of this function is given by,

$$\mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) f_{\mathbf{x}_1, \ldots \mathbf{x}_n}\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) dx_1 \ldots dx_n$$

E.g. $g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) = \mathbf{x}_1 + \cdots + \mathbf{x}_n \implies \mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \sum_{i=1}^{n} \mathbb{E}\left(\mathbf{x}_i\right)$ How do we compute $\mathbb{E}\left(\mathbf{x}_i\right)$? .

.

.

Expected values of multi-variate random variables

- Consider a n-dimensional r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^\top$, and a function $\mathbf{y} = g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)$. The expected value of this function is given by,

$$\mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) f_{\mathbf{x}_1, \ldots \mathbf{x}_n}\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) dx_1 \ldots dx_n$$

E.g. $g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) = \mathbf{x}_1 + \cdots + \mathbf{x}_n \implies \mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \sum_{i=1}^{n} \mathbb{E}\left(\mathbf{x}_i\right)$ How do we compute $\mathbb{E}\left(\mathbf{x}_i\right)$? .

Is $\mathbb{E}\left(\mathbf{x}_i \mathbf{x}_j\right) = \mathbb{E}\left(\mathbf{x}_i\right) \mathbb{E}\left(\mathbf{x}_j\right)$? .

.

Expected values of multi-variate random variables

- Consider a n-dimensional r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^\top$, and a function $\mathbf{y} = g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)$. The expected value of this function is given by,

$$\mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) f_{\mathbf{x}_1, \ldots \mathbf{x}_n}\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) dx_1 \ldots dx_n$$

E.g. $g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right) = \mathbf{x}_1 + \cdots + \mathbf{x}_n \implies \mathbb{E}\left(g\left(\mathbf{x}_1, \ldots \mathbf{x}_n\right)\right) = \sum_{i=1}^{n} \mathbb{E}\left(\mathbf{x}_i\right)$ How do we compute $\mathbb{E}\left(\mathbf{x}_i\right)$? .

Is $\mathbb{E}\left(\mathbf{x}_i \mathbf{x}_j\right) = \mathbb{E}\left(\mathbf{x}_i\right) \mathbb{E}\left(\mathbf{x}_j\right)$? .

What about $\mathbb{E}\left(\Pi_{i=1}^{n} \mathbf{x}_i\right)$? .

Expected values of multi-variate random variables

- **Covariance**. How do two r.v.s $\mathbf{x}_i$ and $\mathbf{x}_j$ co-vary? i.e. how does change in one affect the other?

$$\sigma_{\mathbf{x}_i\mathbf{x}_j} \triangleq \mathbb{E}\Bigg(\left(\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_i\right)\right)\left(\mathbf{x}_j - \mathbb{E}\left(\mathbf{x}_j\right)\right) \Bigg) = \mathbb{E}\left(\mathbf{x}_i\mathbf{x}_j\right) - \mathbb{E}\left(\mathbf{x}_i\right)\mathbb{E}\left(\mathbf{x}_j\right)$$

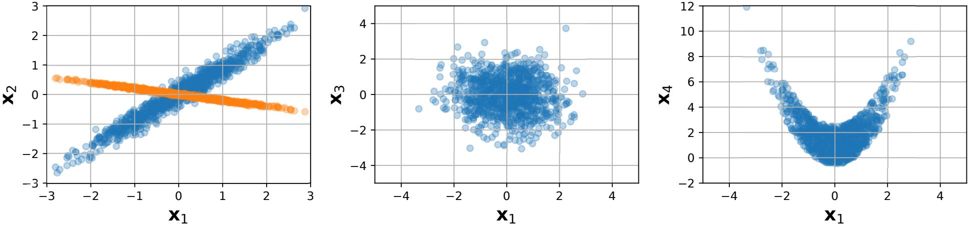
  The size of $\sigma_{\mathbf{x}_i\mathbf{x}_j}$ depends on the variances of the individual random variables $\mathbf{x}_i$ and $\mathbf{x}_j$.

- Positive values for $\sigma_{\mathbf{x}_i\mathbf{x}_j}$ indcates that positive or negative values for $\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_i\right)$ are accompanied by positive or negative values deviations for $\mathbf{x}_j - \mathbb{E}\left(\mathbf{x}_j\right)$, respectively. When for $\sigma_{\mathbf{x}_i\mathbf{x}_j}$ is negative, the signs of $\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_i\right)$ and $\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_j\right)$ are reversed.
- When $\mathbf{x}_i$ and $\mathbf{x}_j$ are independent. $\sigma_{\mathbf{x}_i\mathbf{x}_j} = 0$. Meaning they the signs of $\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_i\right)$ and $\mathbf{x}_i - \mathbb{E}\left(\mathbf{x}_j\right)$ sometimes match and sometimes don't.
- Covariance can be zero even when two r.v.s are dependent. Consider a uniformly distributed r.v. $\mathbf{x}_1$ between -1 and 1, and let $\mathbf{x}_2 = \mathbf{x}_1^2 - \frac{1}{3}$. What is $\sigma_{\mathbf{x}_1\mathbf{x}_2}$?

Expected values of multi-variate random variables

- **Correlation Coefficient**. Normalized covariance and is bound by -1 and +1.

$$\rho_{\mathbf{x}_i\mathbf{x}_j} = \frac{\sigma_{\mathbf{x}_i\mathbf{x}_j}}{\sigma_{\mathbf{x}_i}\sigma_{\mathbf{x}_j}}$$



- $\rho_{\mathbf{x}_i\mathbf{x}_j}$ is 1 or -1 when there is an affine relation between $\mathbf{x}_i$ and $\mathbf{x}_i$, i.e. $\mathbf{x}_i = a\mathbf{x}_j + b$. $\rho_{\mathbf{x}_i\mathbf{x}_j} = 1$ when $a > 0$ and $\rho_{\mathbf{x}_i\mathbf{x}_j} = -1$ when $a < 0$.

Expected values of multi-variate random variables

- Consider a n-dimensional r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^\top$. The mean of $\mathbf{x}$ is given by,

$$\mathbf{m}_\mathbf{x} \triangleq \mathbb{E}\left(\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}\right) = \begin{bmatrix} \mathbb{E}\left(\mathbf{x}_1\right) \\ \vdots \\ \mathbb{E}\left(\mathbf{x}_n\right) \end{bmatrix}$$

- The covaiance matrix of $\mathbf{x}$ is given by,

$$\mathbf{\Sigma}_\mathbf{x} = \mathbb{E}\left(\mathbf{x}^\top \mathbf{x}\right) = \begin{bmatrix} \mathbb{E}\left(\mathbf{x}_1^2\right) & \cdots & \mathbb{E}\left(\mathbf{x}_1\mathbf{x}_n\right) \\ \vdots & \ddots & \vdots \\ \mathbb{E}\left(\mathbf{x}_n\mathbf{x}_1\right) & \cdots & \mathbb{E}\left(\mathbf{x}_n^2\right) \end{bmatrix}$$

All covariance matrices are positive semi-definite.

Multi-variate Gaussian Distributions

- Jointly Gaussian distributions play an important role in many areas of engineering.
- A set of n r.v. $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}^T$ are jointly Gaussian, if their pdf is of the following form.

$$f_{\mathbf{x}}\left(x_1, x_2 \ldots x_n\right) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\mathbf{\Sigma_x}|}} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \mathbf{m_x}\right)^T \mathbf{\Sigma}^{-1} \left(\mathbf{x} - \mathbf{m_x}\right) \right)$$

  The mean $\mathbf{m_x}$ and covariance $\mathbf{\Sigma_x}$ matrix fully characterize the Gaussian distribution.
- Affine transformation of a multivariate Gaussian r.v. results in another Gaussian r.v. with a different mean and covariance. Let, $\mathbf{x} \sim \mathcal{N}\left(\mathbf{m_x}, \mathbf{\Sigma_x}\right)$, and $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$. Then we have,

$$\mathbf{m_y} = \mathbf{A}\mathbf{m_x} + \mathbf{b} \qquad \mathbf{\Sigma_y} = \mathbf{A}\mathbf{\Sigma_x}\mathbf{A}^T$$