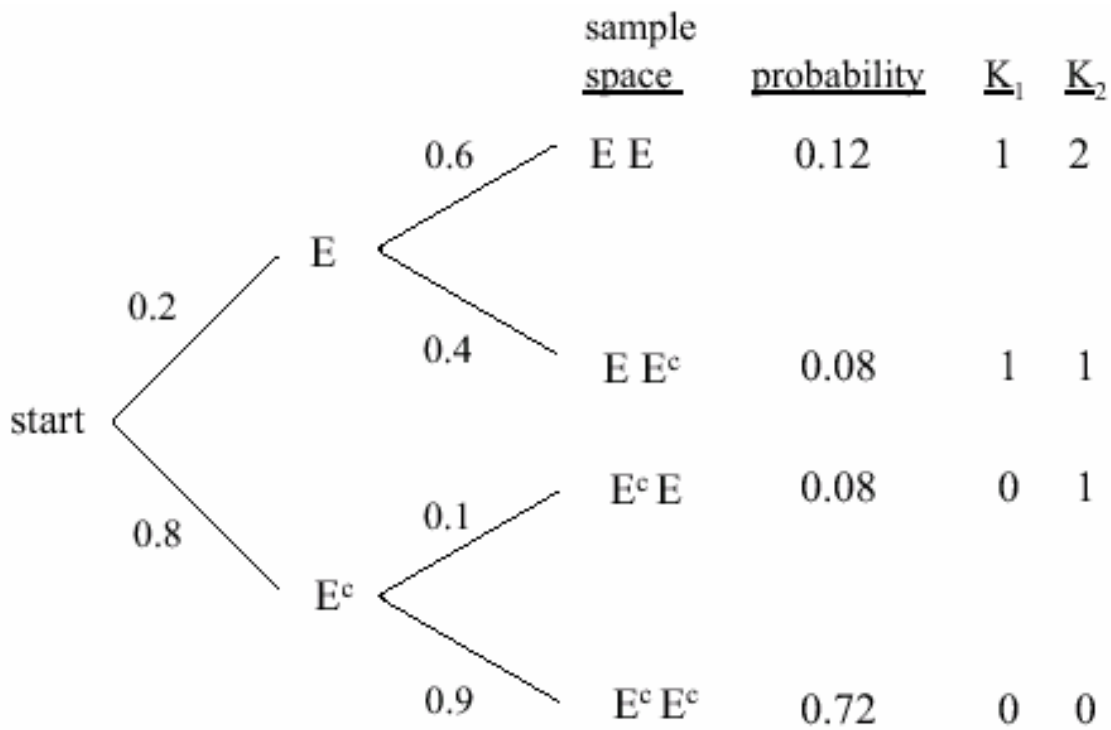
Two Random Variables

Discrete random variables K_1 and K_2.

K_i = Number of errors occurred after observing the ith bit.

Initially Pr[E] =0.2 and Pr[E^c]=0.8



K_1 K_2	probability
1 2	(0.2)(0.6)=0.12
1 1	(0.2)(0.4)=0.08
0 1	(0.8)(0.1)=0.08
0 0	(0.8)(0.9)=0.72

Need joint probability to describe occurrence of two rv's.

© M. Tummala & C. W. Therrien 2004

Joint PMF

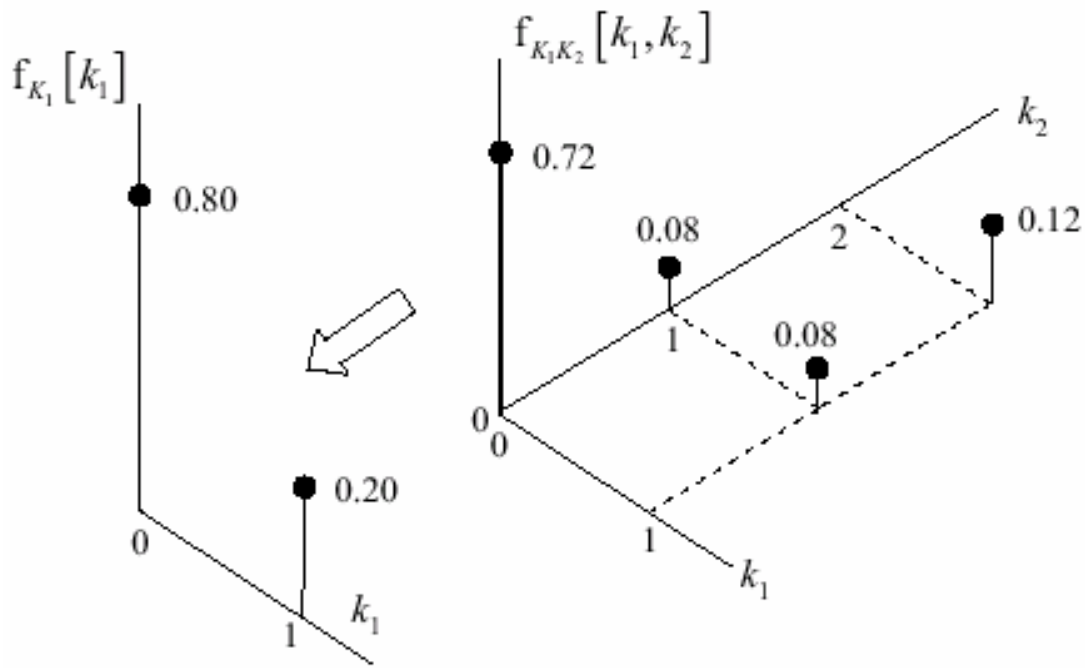
K_1 K_2	probability
1 2	(0.2)(0.6)=0.12
1 1	(0.2)(0.4)=0.08
0 1	(0.8)(0.1)=0.08
0 0	(0.8)(0.9)=0.72

$$\mathrm{f}_{K_1,K_2}[k_1,k_2] = \Pr\left[K_1 = k_1\ ,\ K_2 = k_2\right]$$

$$\sum_{k_1}\sum_{k_2} \mathrm{f}_{K_1,K_2}[k_1,k_2] = 1$$

Marginal PMF:

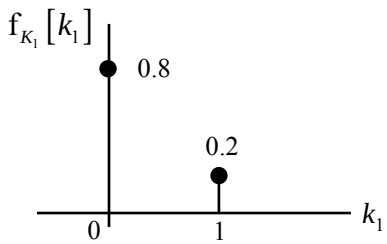
$$\mathrm{f}_{K_1}[k_1] = \sum_{k_2} \mathrm{f}_{K_1,K_2}[k_1,k_2]$$

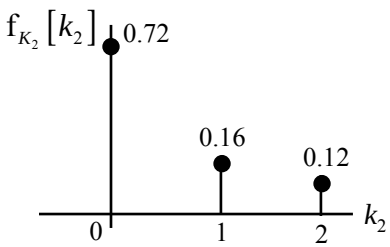


© M. Tummala & C. W. Therrien 2004

Independence for Random Variables

Two rv's are independent iff $\mathrm{f}_{K_1,K_2}[k_1,k_2] = \mathrm{f}_{K_1}[k_1] \cdot \mathrm{f}_{K_2}[k_2]$





		k_2		
		0.72	0.16	0.12
k_1	0.8	0.576	0.128	0.096
	0.2	0.144	0.032	0.024

	k_2		
k_1	0.720	0.080	0
	0	0.080	0.120

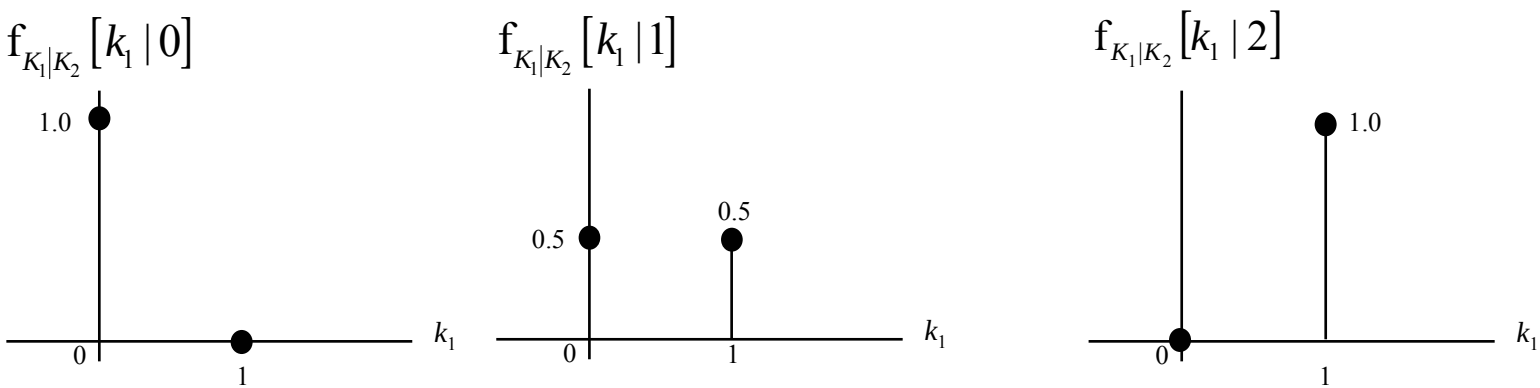
Are these rv's independent?

© M. Tummala & C. W. Therrien 2004

Conditional PMF

Definition: $\mathrm{f}_{K_1|K_2}[k_1 \mid k_2] = \Pr\left[K_1 = k_1 \mid K_2 = k_2\right]$

$$\boxed{\mathrm{f}_{K_1|K_2}[k_1 \mid k_2] = \frac{\mathrm{f}_{K_1K_2}[k_1, k_2]}{\mathrm{f}_{K_2}[k_2]}}$$



<u>Relations:</u>

$$\sum_{k_1} \mathrm{f}_{K_1|K_2}[k_1 \mid k_2] = 1$$

$$\sum_{k_2} \mathrm{f}_{K_1|K_2}[k_1 \mid k_2] = ??!!?$$

© M. Tummala & C. W. Therrien 2004

Bayes' Rule for PMF's

$$\mathrm{f}_{K_1|K_2}[k_1 \mid k_2] = \frac{\mathrm{f}_{K_2|K_1}[k_2 \mid k_1]\mathrm{f}_{K_1}[k_1]}{\mathrm{f}_{K_2}[k_2]}$$

$$= \frac{\mathrm{f}_{K_2|K_1}[k_2 \mid k_1]\mathrm{f}_{K_1}[k_1]}{\sum_{k_1} \mathrm{f}_{K_2|K_1}[k_2 \mid k_1]\mathrm{f}_{K_1}[k_1]}$$

Check this out for the present example!

© M. Tummala & C. W. Therrien 2004

Two Random Variables

Continuous random variables.

Consider two random variables, X_1, X_2

Joint pdf

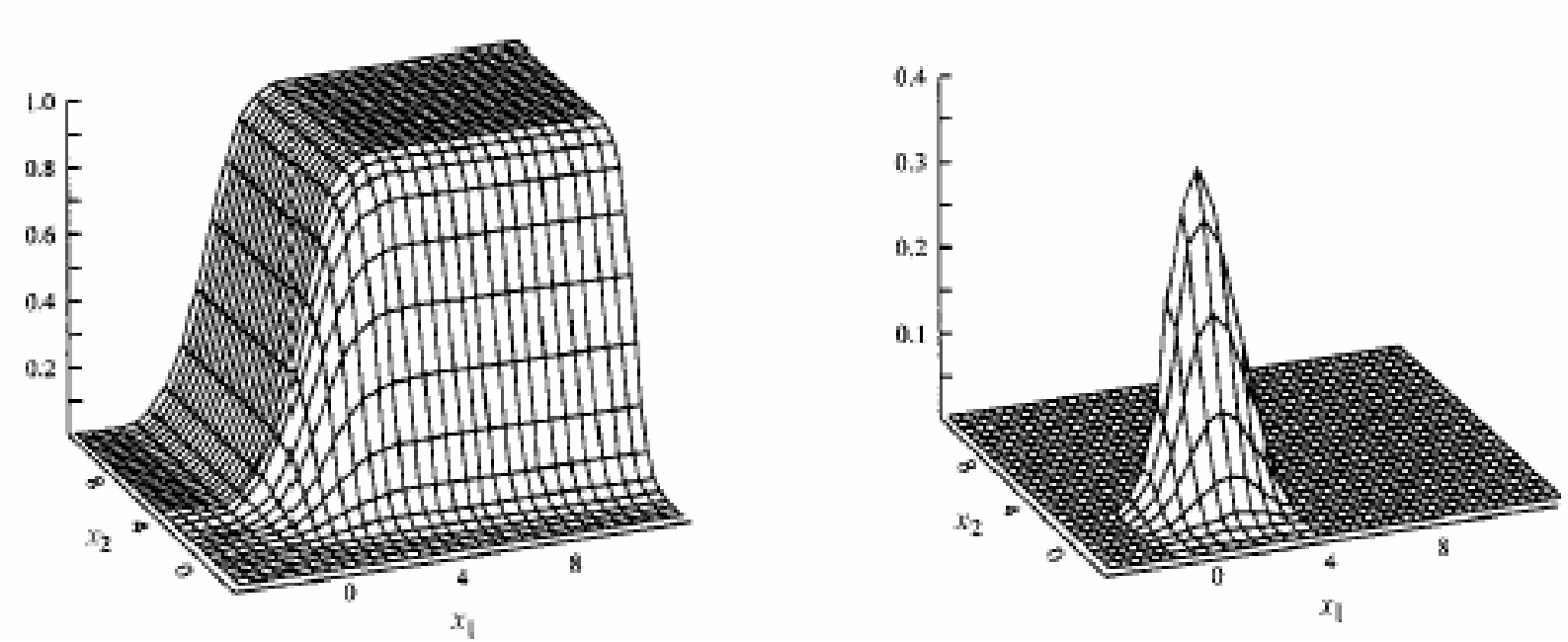
$$f_{X_1,X_2}\left(x_1,x_2\right)=\frac{\partial^2 F_{X_1,X_2}\left(x_1,x_2\right)}{\partial x_1 \partial x_2}$$

Joint cdf

$$\begin{aligned} F_{X_1,X_2}\left(x_1,x_2\right) &= \Pr\left[X_1 \le x_1, X_2 \le x_2\right] \\ &= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1,X_2}\left(z_1,z_2\right) dz_2 dz_1 \end{aligned}$$

© M. Tummala & C. W. Therrien 2004

CDF and PDF



© M. Tummala & C. W. Therrien 2004

Properties of Joint pdf

1. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}\left(x_1,x_2\right)dx_1dx_2 = 1, \qquad f_{X_1,X_2}\left(x_1,x_2\right) \geq 0$

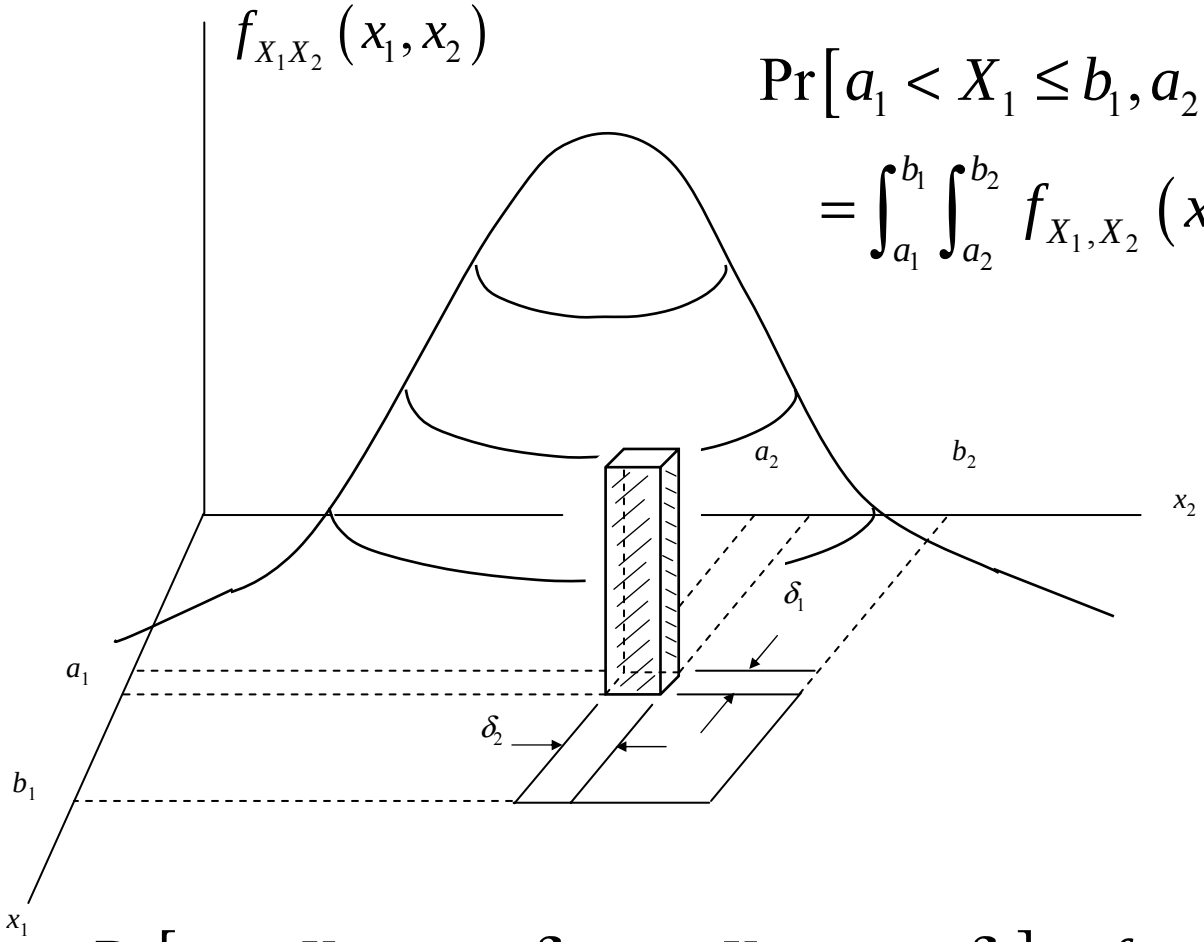
2. $\Pr\left[a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2\right] = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{X_1,X_2}\left(x_1,x_2\right)dx_2dx_1$

3. $\Pr\left[a_1 < X_1 \leq a_1+\delta_1, a_2 < X_2 \leq a_2+\delta_2\right] \cong f_{X_1,X_2}\left(a_1,a_2\right)\delta_1\delta_2$

<u>Note</u>: Equivalent discussion on the same lines exists for discrete random variables.

© M. Tummala & C. W. Therrien 2004

Interpretation of joint PDF as probability:



$$\Pr\left[a_1 < X_1 \le b_1, a_2 < X_2 \le b_2\right] = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{X_1,X_2}(x_1, x_2)\, dx_2 dx_1$$

$$\Pr\left[a_1 < X_1 \le a_1 + \delta_1, a_2 < X_2 \le a_2 + \delta_2\right] \cong f_{X_1,X_2}(a_1, a_2)\,\delta_1\delta_2$$

© M. Tummala & C. W. Therrien 2004

Marginal pdf's

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\, dx_2$$

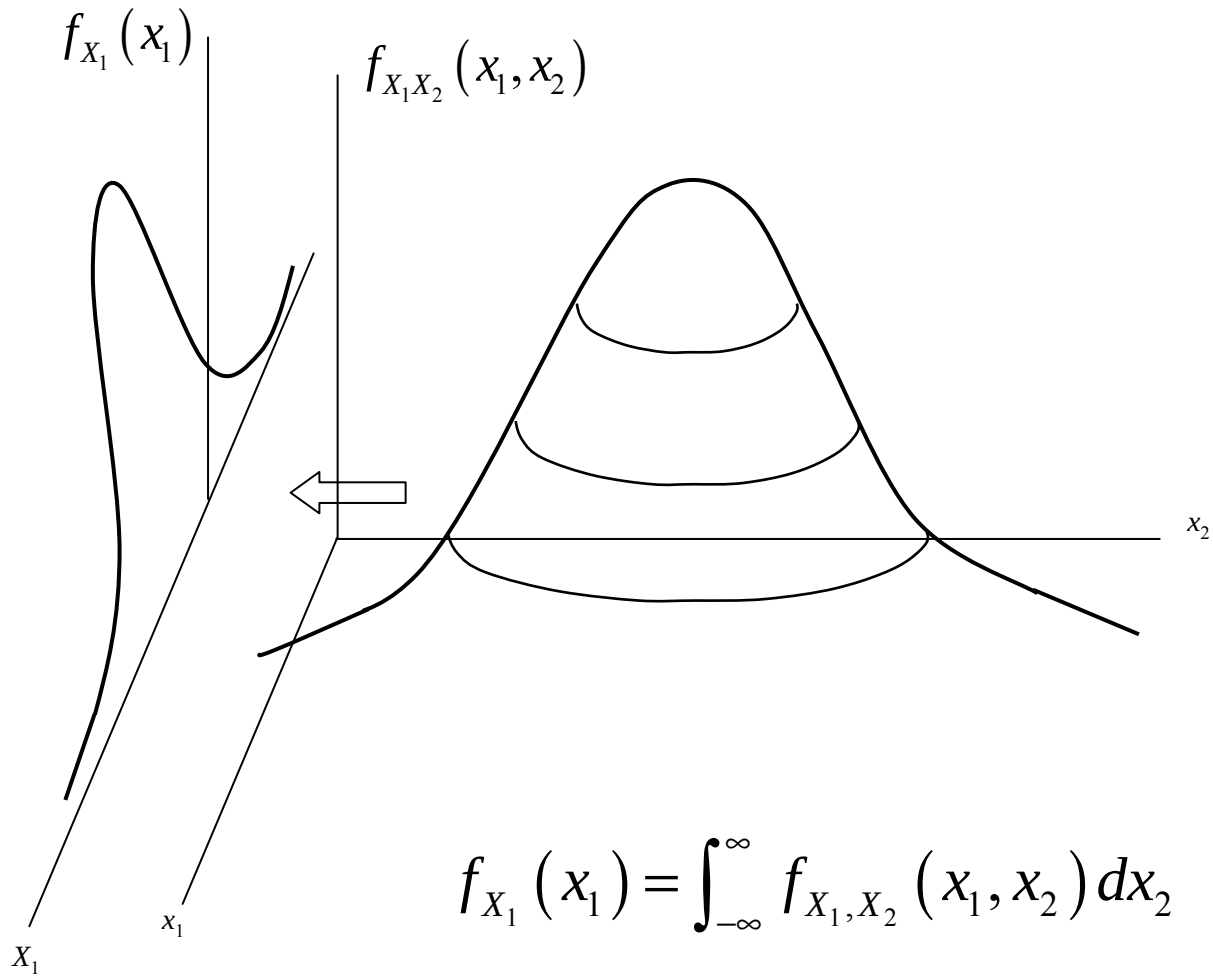
$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\, dx_1$$

Note that:

$$\int_{-\infty}^{\infty} f_{X_1}(x_1)\, dx_1 = \int_{-\infty}^{\infty} f_{X_2}(x_2)\, dx_2 = 1$$

© M. Tummala & C. W. Therrien 2004

Interpretation of marginal pdf as a projection:



$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\, dx_2$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>:

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} ce^{-x_1}e^{-2x_2}, & 0 \le x_1 \le x_2 < \infty \\ 0, & \text{otherwise} \end{cases}$$

(a) Find c

$$1 = c\int_0^\infty \int_0^{x_2} e^{-x_1}e^{-2x_2}\,dx_1\,dx_2 = c\int_0^\infty \left(1 - e^{-x_2}\right) e^{-2x_2}\,dx_2$$

$$1 = c\left[\frac{e^{-2x_2}}{-2} - \frac{e^{-3x_2}}{-3}\right]_0^\infty$$

$$\frac{c}{6} = 1 \quad \Rightarrow \quad c = 6$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>: (continued)

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} 6e^{-x_1}e^{-2x_2}, & 0 \le x_1 \le x_2 < \infty \\ 0, & \text{otherwise} \end{cases}$$

(b) Find $f_{X_1}(x_1)$

$$f_{X_1}(x_1) = 6\int_{x_1}^{\infty} e^{-x_1}e^{-2x_2}\,dx_2 = 3e^{-3x_1}, \qquad 0 \le x_1 < \infty$$

(c) Find $f_{X_2}(x_2)$

$$f_{X_2}(x_2) = 6\int_{0}^{x_2} e^{-x_1}e^{-2x_2}\,dx_1 = 6e^{-2x_2}\left(1 - e^{-x_2}\right)$$

$$0 \le x_2 < \infty$$

© M. Tummala & C. W. Therrien 2004

Properties of Joint cdf

1. $F_{X_1,X_2}(-\infty,-\infty) = F_{X_1,X_2}(-\infty,x_2) = F_{X_1,X_2}(x_1,-\infty) = 0$

   $F_{X_1,X_2}(\infty,\infty) = 1$

2. Marginal cdfs

$$F_{X_1}(x_1) = F_{X_1,X_2}(x_1,\infty)$$

$$F_{X_2}(x_2) = F_{X_1,X_2}(\infty,x_2)$$

3. If $a_1 > a_2$ and $b_1 > b_2$

$$F_{X_1,X_2}(a_1,b_1) \geq F_{X_1,X_2}(a_2,b_2)$$

$\Rightarrow$ monotonically non-decreasing function

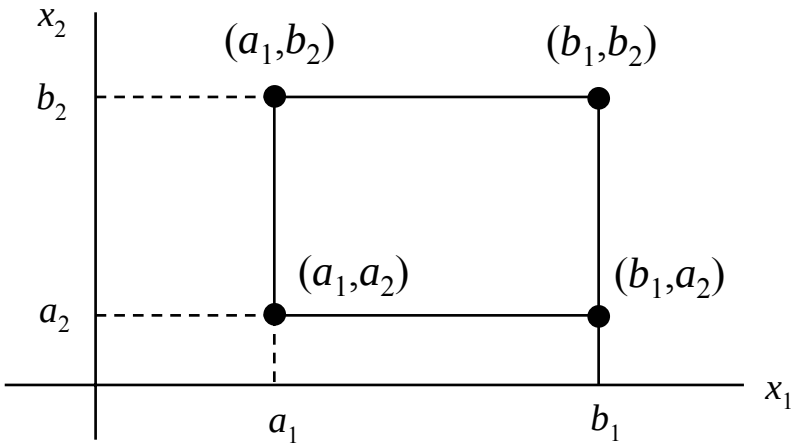
© M. Tummala & C. W. Therrien 2004

Properties of Joint cdf (continued)

4. $\lim_{x_1 \to a^+} F_{X_1X_2}(x_1, x_2) = F_{X_1X_2}(a, x_2)$, continuous from right

$\lim_{x_2 \to b^+} F_{X_1X_2}(x_1, x_2) = F_{X_1X_2}(x_1, b)$, continuous from top

5. $$\Pr\left[a_1 < X_1 \le b_1, a_2 < X_2 \le b_2\right] = F_{X_1,X_2}(b_1, b_2) - F_{X_1,X_2}(a_1, b_2) - F_{X_1,X_2}(b_1, a_2) + F_{X_1,X_2}(a_1, a_2)$$



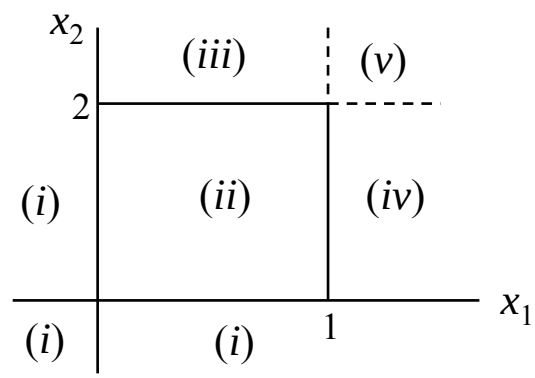
© M. Tummala & C. W. Therrien 2004

<u>Example</u>:

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} 0.5, & 0 \le x_1 \le 1,\ 0 \le x_2 < 2 \\ 0, & \text{otherwise} \end{cases}$$

Find the joint cdf.

$$F_{X_1,X_2}(x_1, x_2) = \Pr[X_1 \le x_1, X_2 \le x_2] = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1,X_2}(z_1, z_2)\, dz_2 dz_1$$



<u>*Case i*</u> $x_1 < 0$ or $x_2 < 0$ or both $x_1, x_2 < 0$

$$F_{X_1,X_2}(x_1, x_2) = 0$$

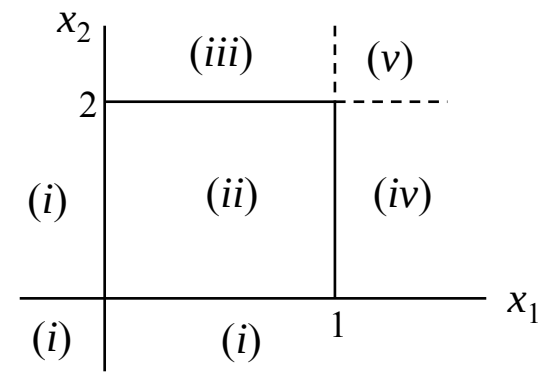
© M. Tummala & C. W. Therrien 2004

Case ii $0 \le x_1 \le 1,\ 0 \le x_2 \le 2$

$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{x_1}\int_0^{x_2} dz_2 dz_1 = \frac{1}{2}x_1 x_2$$

Case iii $0 \le x_1 \le 1,\ x_2 > 2$

$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{x_1}\int_0^{2} dz_2 dz_1 = x_1$$



Case iv $x_1 > 1,\ 0 \le x_2 \le 2$

$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{1}\int_0^{x_2} dz_2 dz_1 = \frac{1}{2}x_2$$

Case v $x_1 > 1, x_2 > 2$

$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{1}\int_0^{2} dz_2 dz_1 = 1$$

© M. Tummala & C. W. Therrien 2004

Independence of Random Variables

Two random variables X_1 and X_2 are said to be *independent* if

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$$
$$\text{for all } x_1, x_2$$

or

$$F_{X_1,X_2}(x_1,x_2) = F_{X_1}(x_1)\, F_{X_2}(x_2)$$
$$\text{for all } x_1, x_2$$

© M. Tummala & C. W. Therrien 2004

Example:

$$f_{X_1X_2}(x_1,x_2)=\begin{cases}6e^{-x_1}e^{-2x_2}, & 0\le x_1\le x_2<\infty\\ 0, & \text{otherwise}\end{cases}$$

$$f_{X_1}(x_1)=6\int_{x_1}^{\infty}e^{-x_1}e^{-2x_2}\,dx_2=3e^{-3x_1},\quad 0\le x_1<\infty$$

$$f_{X_2}(x_2)=6\int_{0}^{x_2}e^{-x_1}e^{-2x_2}\,dx_1=6e^{-2x_2}\left(1-e^{-x_2}\right),\quad 0\le x_2<\infty$$

(a) Are X_1, X_2 independent?

$$f_{X_1X_2}(x_1,x_2)\overset{?}{=}f_{X_1}(x_1)f_{X_2}(x_2)$$

$$6e^{-x_1}e^{-2x_2}\neq 18e^{-3x_1}e^{-2x_2}\left(1-e^{-x_2}\right)\qquad\Leftarrow\text{not independent}$$

© M. Tummala & C. W. Therrien 2004

Example:

$$F_{X_1,X_2}(x_1,x_2)=\begin{cases}1-e^{-ax_1}-e^{-bx_2}+e^{-(ax_1+bx_2)} & x_1\geq 0, x_2\geq 0\\ 0 & \text{otherwise}\end{cases}$$

(a) Determine the marginal cdfs

$$F_{X_1}(x_1)=F_{X_1X_2}(x_1,\infty)=1-e^{-ax_1} \qquad x_1\geq 0$$

$$F_{X_2}(x_2)=F_{X_1X_2}(\infty,x_2)=1-e^{-bx_2} \qquad x_2\geq 0$$

(b) Obtain the marginal pdfs from (a)

$$f_{X_1}(x_1)=\frac{dF_{X_1}(x_1)}{dx_1}=ae^{-ax_1} \qquad x_1\geq 0$$

$$f_{X_2}(x_2)=\frac{dF_{X_2}(x_2)}{dx_2}=be^{-bx_2} \qquad x_2\geq 0$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>: (continued)

(c) Are X_1, X_2 independent? Use cdfs

$$F_{X_1X_2}\left(x_1, x_2\right) \stackrel{?}{=} F_{X_1}\left(x_1\right) F_{X_2}\left(x_2\right)$$

$$1 - e^{-ax_1} - e^{-bx_2} + e^{-\left(ax_1 + bx_2\right)} = \left(1 - e^{-ax_1}\right)\left(1 - e^{-bx_2}\right)$$

Yes, they are independent

(d) Determine their independence using pdfs

$$f_{X_1X_2}\left(x_1, x_2\right) = \frac{\partial F_{X_1X_2}\left(x_1 x_2\right)}{\partial x_1 \partial x_2} = abe^{-ax_1}e^{-bx_2} \qquad x_1 \geq 0, x_2 \geq 0$$

$$f_{X_1X_2}\left(x_1, x_2\right) \stackrel{?}{=} f_{X_1}\left(x_1\right) f_{X_2}\left(x_2\right)$$

$$abe^{-ax_1}e^{-bx_2} = ae^{-ax_1}be^{-bx_2}$$ Yes, they are independent

© M. Tummala & C. W. Therrien 2004

Conditional pdf

Given two random variables X_1 and X_2, we can write the following conditional density functions:

$$f_{X_1|X_2}\left(x_1 \middle| x_2\right) = \frac{f_{X_1,X_2}\left(x_1, x_2\right)}{f_{X_2}\left(x_2\right)}$$

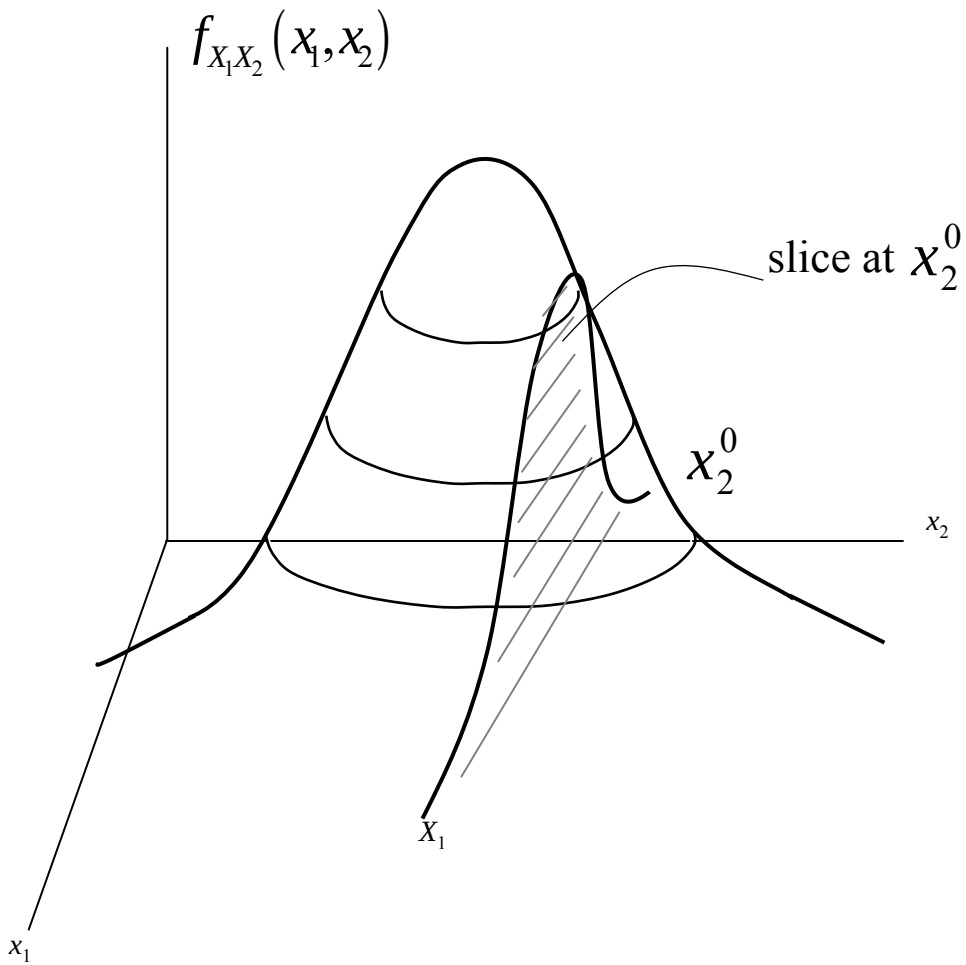
$$f_{X_2|X_1}\left(x_2 \middle| x_1\right) = \frac{f_{X_1,X_2}\left(x_1, x_2\right)}{f_{X_1}\left(x_1\right)}$$

Note that:

$$\int_{-\infty}^{\infty} f_{X_1|X_2}\left(x_1 \middle| x_2\right) dx_1 = 1 \quad \text{while} \quad \int_{-\infty}^{\infty} f_{X_1|X_2}\left(x_1 \middle| x_2\right) dx_2 = ???$$

© M. Tummala & C. W. Therrien 2004

Interpretation of conditional pdf as a slice through the joint pdf:



$$f_{X_1|X_2}\left(x_1 \middle| x_2^0\right) = \frac{f_{X_1X_2}\left(x_1, x_2^0\right)}{f_{X_2}\left(x_2^0\right)}$$

© M. Tummala & C. W. Therrien 2004

Example:

$$f_{X_1X_2}(x_1,x_2)=\begin{cases}6e^{-x_1}e^{-2x_2}, & 0\le x_1\le x_2<\infty\\ 0, & \text{otherwise}\end{cases}$$

$$f_{X_1}(x_1)=6\int_{x_1}^{\infty}e^{-x_1}e^{-2x_2}dx_2=3e^{-3x_1},\quad 0\le x_1<\infty$$

$$f_{X_2}(x_2)=6\int_{0}^{x_2}e^{-x_1}e^{-2x_2}dx_1=6e^{-2x_2}\left(1-e^{-x_2}\right),\quad 0\le x_2<\infty$$

Find conditional pdfs

$$f_{X_1|X_2}(x_1\mid x_2)=\frac{f_{X_1X_2}(x_1,x_2)}{f_{X_2}(x_2)}=\frac{6e^{-x_1}e^{-2x_2}}{6e^{-2x_2}(1-e^{-x_2})}=\frac{e^{-x_1}}{1-e^{-x_2}},\quad 0\le x_1\le x_2$$

$$f_{X_2|X_1}(x_2\mid x_1)=\frac{f_{X_1X_2}(x_1,x_2)}{f_{X_1}(x_1)}=\frac{6e^{-x_1}e^{-2x_2}}{3e^{-3x_1}}=2e^{2x_1}e^{-2x_2},\quad x_1\le x_2<\infty$$

© M. Tummala & C. W. Therrien 2004

Bayes' Rule for densities

Given two random variables X_1 and X_2, we can write the following:

$$f_{X_1|X_2}(x_1 \mid x_2) = \frac{f_{X_2|X_1}(x_2 \mid x_1)\, f_{X_1}(x_1)}{f_{X_2}(x_2)}$$

Now since:

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2)\,dx_1 = \int_{-\infty}^{\infty} f_{X_2|X_1}(x_2 \mid x_1)\, f_{X_1}(x_1)\,dx_1$$

we have:

$$\boxed{f_{X_1|X_2}(x_1 \mid x_2) = \frac{f_{X_2|X_1}(x_2 \mid x_1)\, f_{X_1}(x_1)}{\int_{-\infty}^{\infty} f_{X_2|X_1}(x_2 \mid x_1)\, f_{X_1}(x_1)\,dx_1}}$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>:

Given a conditional pdf and the marginal pdfs:

$$f_{X_1}(x_1 \mid x_2) = \frac{e^{-x_1}}{1-e^{-x_2}}, \;\; 0 \le x_1 \le x_2$$

$$f_{X_1}(x_1) = 6\int_{x_1}^{\infty} e^{-x_1} e^{-2x_2}\, dx_2 = 3e^{-3x_1}, \quad 0 \le x_1 < \infty$$

$$f_{X_2}(x_2) = 6\int_{0}^{x_2} e^{-x_1} e^{-2x_2}\, dx_1 = 6e^{-2x_2}\left(1-e^{-x_2}\right), \;\; 0 \le x_2 < \infty$$

Find the inverse conditional pdf using the Bayes' rule:

$$f_{X_2|X_1}(x_2 \mid x_1) = \frac{f_{X_1|X_2}(x_1 \mid x_2)\, f_{X_2}(x_2)}{f_{X_1}(x_1)}$$

$$= \frac{6e^{-x_1} e^{-2x_2}\left(1-e^{-x_2}\right)}{3e^{-3x_1}(1-e^{-x_2})} = 2e^{2x_1} e^{-2x_2}, \;\; x_1 \le x_2 < \infty$$

© M. Tummala & C. W. Therrien 2004

1st and 2nd Order Moments for Two Random Variables

<u>mean</u>

$$m_i = E\left[X_i\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i \; f_{X_1X_2}\left(x_1, x_2\right) dx_1 dx_2, \qquad i = 1,\ 2$$

<u>variance</u>

$$\sigma_i^2 = \mathrm{Var}\left[X_i\right] = E\left[(X_i - m_i)^2\right] \qquad i = 1,\ 2$$

<u>correlation</u>

$$r_{ij} = E\left[X_i X_j\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i x_j \, f_{X_iX_j}\left(x_i, x_j\right) dx_i dx_j \qquad i, j = 1,2$$

<u>covariance</u>

$$c_{ij} = \mathrm{Cov}\left[X_i, X_j\right] = E\left[(X_i - m_i)(X_j - m_j)\right] \qquad i, j = 1,2$$

© M. Tummala & C. W. Therrien 2004

Correlation/Covariance Relations

$$c_{ij} = r_{ij} - m_i m_j$$

$$\mathrm{Cov}\left[X_i, X_j\right] = E\left[X_i X_j\right] - E\left[X_i\right] \cdot E\left[X_j\right]$$

➡ If $\mathrm{Cov}[X_i, X_j] = 0$ then X_i and X_j are said to be <u>uncorrelated</u>.

Note that in this case

$$E\left[X_1 X_2\right] = E\left[X_1\right] E\left[X_2\right]$$

(If $E[X_i X_j] = 0$, the random variables are said to be *orthogonal*.)

© M. Tummala & C. W. Therrien 2004

More Relations for Two Random Variables

<u>Independent</u> random variables are <u>uncorrelated</u>:

<u>Proof</u>:

$$E\left[X_1 X_2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X_1 X_2}\left(x_1, x_2\right) dx_1 dx_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X_1}\left(x_1\right) f_{X_2}\left(x_2\right) dx_1 dx_2$$

$$\left(\text{since } X_1, X_2 \text{ are independent}\right)$$

$$E\left[X_1 X_2\right] = \int_{-\infty}^{\infty} x_1 f_{X_1}\left(x_1\right) dx_1 \int_{-\infty}^{\infty} x_2 f_{X_2}\left(x_2\right) dx_2 = E\left[X_1\right] E\left[X_2\right]$$

$\therefore$ X_1 and X_2 are *uncorrelated*

<u>vice versa is not true</u> (except for Gaussian random variables)

© M. Tummala & C. W. Therrien 2004

Summary of Correlation Relations

- X_1 and X_2 are <u>uncorrelated</u> if Cov $[X_1, X_2] = 0$
- X_1 and X_2 are <u>orthogonal</u> if

$$E\left[X_1 X_2\right] = 0$$

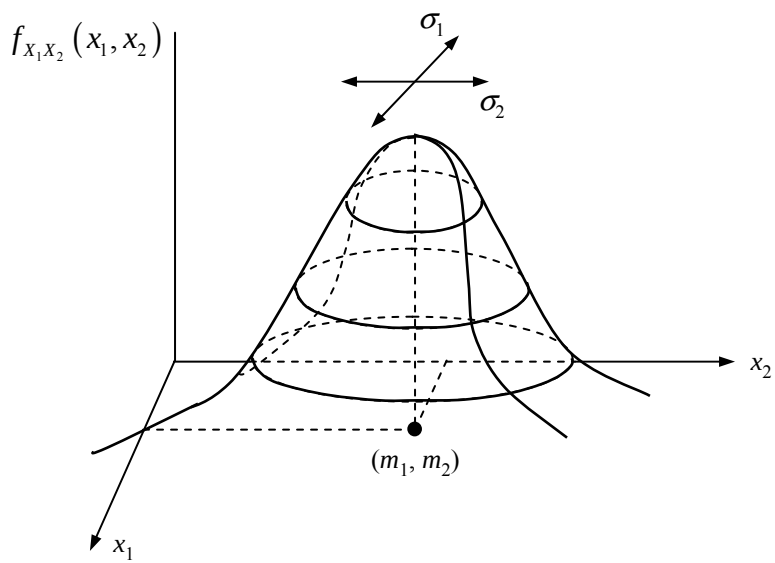
- The <u>correlation coefficient</u> of X_1 and X_2 is defined as

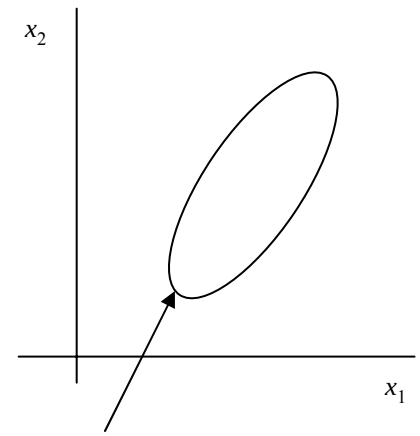
$$\rho = \frac{E\left[\left(X_1 - m_1\right)\left(X_2 - m_2\right)\right]}{\sigma_{X_1}\sigma_{X_2}} = \frac{\mathrm{Cov}\left[X_1, X_2\right]}{\sigma_{X_1}\sigma_{X_2}}$$

$$-1 \le \rho \le 1$$

© M. Tummala & C. W. Therrien 2004

Bivariate Gaussian PDF (Joint Gaussian density function)





contour

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$

$$\exp -\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right]$$

© M. Tummala & C. W. Therrien 2004

Bivariate Gaussian PDF, cont'd.

Marginal densities

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp - \left[\frac{(x_i - m_i)^2}{2\sigma_i^2} \right] \qquad i = 1, 2$$

Conditional density

$$f_{X_1|X_2}(x_1 \mid x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \left[\frac{(x_1 - \mu(x_2))^2}{2\sigma^2} \right]$$

where

$$\sigma^2 = \sigma_1^2 (1 - \rho^2) \quad \text{and} \quad \mu(x_2) = m_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - m_2)$$

© M. Tummala & C. W. Therrien 2004

Expectation of Two Random Variables (Random Vectors)

Let us use the compact notation

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad 2\times 1 \text{ vector,} \qquad f_{\mathbf{X}}(\boldsymbol{x}), F_{\mathbf{X}}(\boldsymbol{x})$$

$$E\left[\gamma(\mathbf{X})\right] = \int_{-\infty}^{\infty} \gamma(\boldsymbol{x})\, f_{\mathbf{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

$$E\left[\gamma(X_1, X_2)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(x_1, x_2)\, f_{X_1X_2}(x_1, x_2)\, dx_1 dx_2$$

Mean of X (**mean vector**)

$$E[\mathbf{X}] = E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right] = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \mathbf{m}_X \qquad 2\times 1 \text{ vector}$$

© M. Tummala & C. W. Therrien 2004

Second moment (**correlation matrix**)

$$\mathbf{R}_X = E\left[\mathbf{X}\mathbf{X}^T\right] = E\left[\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix}\right]$$

$$= E\begin{bmatrix} X_1X_1 & X_1X_2 \\ X_2X_1 & X_2X_2 \end{bmatrix}$$

$$= \begin{bmatrix} E\left[X_1^2\right] & E\left[X_1X_2\right] \\ E\left[X_2X_1\right] & E\left[X_2^2\right] \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

When the correlation of X_1 and X_2 is defined as

$$r_{ij} = E\left[X_iX_j\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i x_j \; f_{X_1X_2}\left(x_1, x_2\right) dx_1 dx_2, \qquad i = 1,2;\; j = 1,2$$

© M. Tummala & C. W. Therrien 2004

Second central moment (**covariance matrix**)

$$\mathbf{C}_X = E\left[\left(\mathbf{X}-\mathbf{m}_X\right)\left(\mathbf{X}-\mathbf{m}_X\right)^T\right]$$

$$= \begin{bmatrix} E\left[\left(X_1-m_1\right)^2\right] & E\left[\left(X_1-m_1\right)\left(X_2-m_2\right)\right] \\ E\left[\left(X_2-m_2\right)\left(X_1-m_1\right)\right] & E\left[\left(X_2-m_2\right)^2\right] \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

where the covariance of X_1 and X_2 is defined as

$$c_{ij} = \mathrm{Cov}\left[X_i, X_j\right] = E\left[\left(X_i-m_i\right)\left(X_j-m_j\right)\right]$$

Notice that $\quad c_{ii} = \sigma_{X_i}^2 = \mathrm{Var}\left[X_i\right]$

© M. Tummala & C. W. Therrien 2004

An important relation:

$$\mathbf{C}_X = \mathbf{R}_X - \mathbf{m}_X \mathbf{m}_X^T$$

Proof:

$$\begin{aligned}
\mathbf{C}_X &= E\left[\left(\mathbf{X}-\mathbf{m}_X\right)\left(\mathbf{X}-\mathbf{m}_X\right)^T\right] \\
&= E\left[\mathbf{X}\mathbf{X}^T\right] - E\left[\mathbf{X}\right]\mathbf{m}_X^T - \mathbf{m}_X E\left[\mathbf{X}^T\right] + \mathbf{m}_X \mathbf{m}_X^T \\
&= \mathbf{R}_X - \mathbf{m}_X \mathbf{m}_X^T - \mathbf{m}_X \mathbf{m}_X^T + \mathbf{m}_X \mathbf{m}_X^T \\
&= \mathbf{R}_X - \mathbf{m}_X \mathbf{m}_X^T
\end{aligned}$$

In component form: $c_{ij} = r_{ij} - m_i m_j$

(which we have seen before)

© M. Tummala & C. W. Therrien 2004

Bivariate Gaussian Random Variables X_1, X_2

Let X_1 and X_2 be jointly Gaussian

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \mathbf{m}_X = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \quad \mathbf{C}_X = \begin{bmatrix} \sigma_1^2 & \operatorname{cov}(X_1, X_2) \\ \operatorname{cov}(X_1, X_2) & \sigma_2^2 \end{bmatrix}$$

Gaussian pdf for an $N \times 1$ vector

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{(2\pi)^{N/2} \left| C_X \right|^{1/2}} e^{-1/2(\mathbf{x}-\mathbf{m}_X)^T C_X^{-1} (\mathbf{x}-\mathbf{m}_X)}$$

For $N = 2$ and by letting $\sigma_1^2 = \sigma_2^2 = \sigma^2$,

$$C_X = \begin{bmatrix} \sigma^2 & \operatorname{cov}(X_1, X_2) \\ \operatorname{cov}(X_1, X_2) & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

where $\rho = \operatorname{cov}(X_1, X_2) / \sigma^2$

© M. Tummala & C. W. Therrien 2004

Examples:

(a) Let $N = 2$, $\sigma^2 = 4$, $\rho = 0.8$, $\mathbf{m}_X = 0$

$$\left|\mathbf{C}_X\right|^{1/2} = \sigma^2\sqrt{1-\rho^2} = 2.4,$$

$$\mathbf{C}_X^{-1} = \frac{1}{\sigma^2\left(1-\rho^2\right)}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} = \frac{1}{1.44}\begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{C}_X^{-1}\mathbf{x} = \frac{1}{2.88}\begin{bmatrix} x_1 - 0.8x_2 & -0.8x_1 + x_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

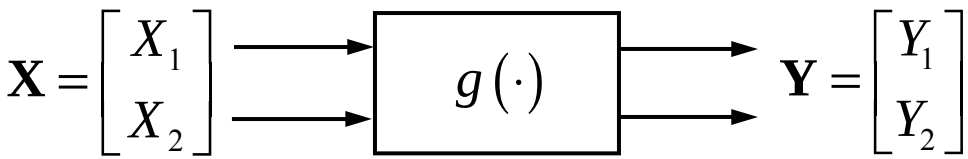
$$f_{X_1X_2}\left(x_1, x_2\right) = \frac{1}{4.8\pi}e^{-\left(x_1^2 - 1.6x_1x_2 + x_2^2\right)/2.88}$$

(b) Let X_1 and X_2 be independent: $\rho = 0$; also $\mathbf{m}_X = 0$, $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$f_{X_1X_2}\left(x_1, x_2\right) = \frac{1}{2\pi\sigma^2}e^{-\left(x_1^2 + x_2^2\right)/2\sigma^2} = \frac{1}{\sqrt{2\pi}\,\sigma}e^{-x_1^2/2\sigma^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma}e^{-x_2^2/2\sigma^2}$$

© M. Tummala & C. W. Therrien 2004

Functions of Two r.v.s



$$\mathbf{Y} = g(\mathbf{X}) = g(X_1, X_2)$$

$$f_{\mathbf{Y}}(\mathbf{y}) = \left| J(\mathbf{y}) \right| f_{\mathbf{X}}(x) \Big|_{\mathbf{x}=g^{-1}(\mathbf{y})}$$

Two r.v.s

$$\left| J(y_1, y_2) \right| = \left\| \begin{bmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{bmatrix} \right\| = \frac{1}{\left| J(x_1, x_2) \right|}$$

Single r.v.

$$\left| J(y) \right| = \left| \frac{dx}{dy} \right|$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>:

X_1, X_2 are zero-mean, independent Gaussian r.v.s with a variance σ^2

$$f_{X_1X_2}\left(x_1, x_2\right) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x_1^2/2\sigma^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x_2^2/2\sigma^2}$$

$$= \frac{1}{2\pi\sigma^2} e^{-\left(x_1^2+x_2^2\right)/2\sigma^2}, \qquad -\infty < x_1, x_2 < \infty$$

Output random variables (cartesian to polar coordinates)

$$\begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \tan^{-1}\left(\dfrac{x_2}{x_1}\right) \end{bmatrix}, \qquad \begin{array}{l} r \geq, \text{ magnitude} \\ \theta = \left[0, 2\pi\right], \text{ angle} \end{array}$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>: (continued)

The inverse mapping is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}$$

$$\left|J\left(r,\theta\right)\right| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r$$

The joint density function of the output vector

$$f_{R\theta}\left(r,\theta\right) = r \cdot \frac{1}{2\pi\sigma^2} \cdot e^{-r^2/2\sigma^2} \qquad r \geq 0, \quad 0 \leq \theta \leq 2\pi.$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>: (continued)

The marginal density function of R

$$f_R(r) = \int_0^{2\pi} f_{R\theta}(r,\theta)\,d\theta = \frac{r}{2\pi\sigma^2} e^{-r^2/2\sigma^2} \int_0^{2\pi} d\theta$$

$$= \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \qquad r \geq 0 \qquad \text{Rayleigh r.v.}$$

The marginal density function of θ

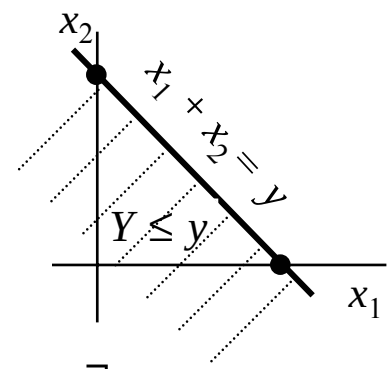
$$f_\theta(\theta) = \int_0^{\infty} f_{R\theta}(r,\theta)\,dr = \frac{1}{2\pi} \int_0^{\infty} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}\,dr$$

$$= \frac{1}{2\pi} \qquad 0 \leq \theta \leq 2\pi \qquad \text{Uniform r.v.}$$

© M. Tummala & C. W. Therrien 2004

<u>Sum of two random variables</u> $Y = X_1 + X_2$

Determine the pdf of Y from its cdf



$$F_Y(y) = \Pr[Y \le y] = \Pr[X_1 + X_2 \le y]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{y-x_1} f_{X_1X_2}(x_1, x_2)\, dx_2 dx_1$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \int_{-\infty}^{\infty} \left[\frac{d}{dy} \int_{-\infty}^{y-x_1} f_{X_1X_2}(x_1, x_2)\, dx_2 \right] dx_1$$

$$= \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, y - x_1)\, dx_1$$

$$= \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(y - x_1)\, dx_1 \qquad \text{if } X_1, X_2 \text{ are independent}$$

Sum of two independent random variables: The pdf of the sum is a <u>convolution</u> of the component pdfs.

© M. Tummala & C. W. Therrien 2004

Example:

X_1, X_2 are lifetimes of two light bulbs that are used sequentially.

The combined lifetime of two light bulbs is : $Y = X_1 + X_2$.

The pdf of X_i, $i = 1, 2$: $$f_{X_i}(x_i) = \lambda e^{-\lambda x_i} \qquad x_i \geq 0$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(y - x_1)\, dx_1$$

Substitute for the pdfs of X_1 and X_2:

$$f_Y(y) = \int_0^y \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda(y - x_1)}\, dx_1$$

$$= \lambda^2 e^{-\lambda y} \int_0^y dx_1 = \lambda^2 y e^{-\lambda y}, \qquad y \geq 0$$

(Y is a 2-Erlang r.v.)

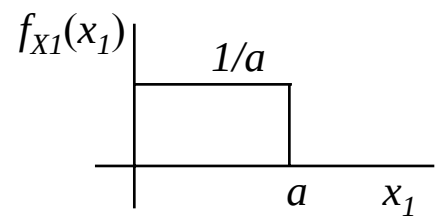
© M. Tummala & C. W. Therrien 2004

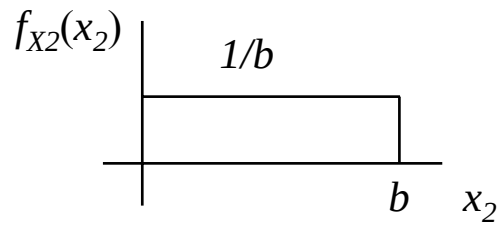
<u>Example</u>:

Consider the sum of two random variables $Y = X_1 + X_2$. The pdf $f_{X_1}(x_1)$ is uniform $[0, a]$ and $f_{X_2}(x_2)$ is uniform $[0, b]$; $0 < a < b$.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(y - x_1)\, dx_1$$

<u>*Case 1*</u> $y < 0, f_Y(y) = 0$





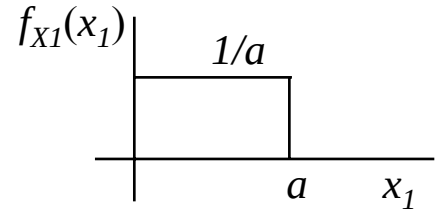
<u>*Case 2*</u> $0 \le y < a$

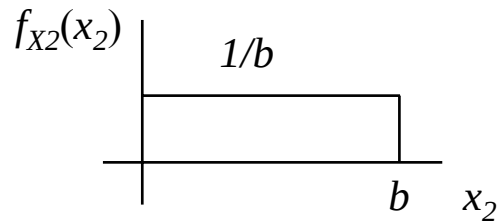
$$f_Y(y) = \int_0^y \frac{1}{a} \cdot \frac{1}{b}\, dx_1 = \frac{1}{ab}\, y$$

© M. Tummala & C. W. Therrien 2004

<u>*Case 3*</u> $a \le y < b$

$$f_Y(y) = \int_0^a \frac{1}{a} \cdot \frac{1}{b} dx_1 = \frac{1}{b}$$



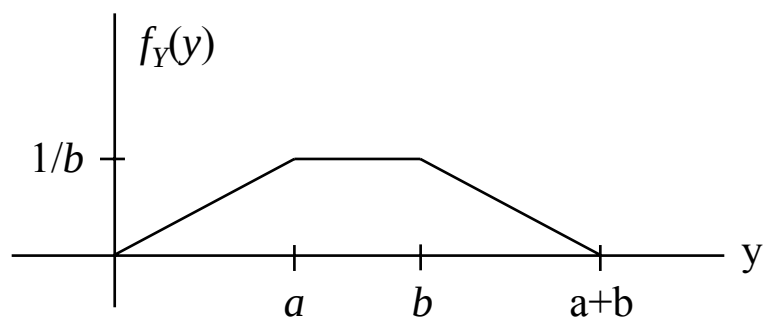


<u>*Case 4*</u> $b \le y \le a + b$

$$f_Y(y) = \int_{y-b}^{a} \frac{1}{a} \cdot \frac{1}{b} dx_1 = \frac{1}{ab} x_1 \Big|_{y-b}^{a} = \frac{1}{ab}\left[a + b - y\right]$$

<u>*Case 5*</u> $y > a + b$

$$f_Y(y) = 0$$



© M. Tummala & C. W. Therrien 2004

Sums of r.v.s

Let $X_1, X_2, \ldots, X_n$ be n random variables. Let their sum be

$$Y = X_1 + X_2 + \cdots + X_n$$

Mean of Y

$$E\left[X_1 + \cdots + X_n\right] = E\left[X_1\right] + E\left[X_2\right] + \cdots + E\left[X_n\right]$$

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E\left[X_i\right]$$

© M. Tummala & C. W. Therrien 2004

Sums of r.v.s (cont'd)

Variance of Y (let $n = 2$ for simplicity)

$$\begin{aligned}
\operatorname{var}(X_1 + X_2) &= E\left[\left(X_1 + X_2 - E[X_1] - E[X_2]\right)^2\right] \\
&= E\left[\left((X_1 - m_1) + (X_2 - m_2)\right)^2\right] \\
&= E\left[(X_1 - m_1)^2\right] + E\left[(X_2 - m_2)^2\right] + 2E\left[(X_1 - m_1)(X_2 - m_2)\right] \\
&= \operatorname{var}(X_1) + \operatorname{var}(X_2) + 2\operatorname{cov}(X_1, X_2)
\end{aligned}$$

Generalizing, we have

$$\operatorname{var}(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} \operatorname{var}(X_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \operatorname{cov}(X_i, X_j)$$

© M. Tummala & C. W. Therrien 2004

Variance for Uncorrelated Random Variables

$$Y = X_1 + X_2 + \cdots + X_n$$

If X_i and X_j are uncorrelated ($i \neq j$), the covariance is zero:

$$E\left[\left(X_i - m_X\right)\left(X_j - m_X\right)\right] = \operatorname{cov}\left(X_i, X_j\right) = 0$$

Therefore

$$\boxed{\operatorname{var}\left(X_1 + X_2 + \cdots + X_n\right) = \sum_{i=1}^{n} \operatorname{var}\left(X_i\right)}$$

for *uncorrelated* r.v.'s

- Independent random variables are uncorrelated.

© M. Tummala & C. W. Therrien 2004

IID Random Variables

Independent and identically-distributed (IID) r.v.s are mutually independent and all have the same pdf/cdf.

Consider that $X_1, X_2, \ldots, X_n$ are IID, and each has a mean m_X and variance σ_X^2

Then

$$E\left[X_1 + \cdots + X_n\right] = \sum_{i=1}^{n} E\left[X_i\right] = n\, m_X$$

Since X_i and X_j are independent ($i \neq j$), the covariance is zero:

$$E\left[\left(X_i - m_X\right)\left(X_j - m_X\right)\right] = \operatorname{cov}\left(X_i, X_j\right) = 0$$

As a result

$$\operatorname{var}\left(X_1 + X_2 + \cdots + X_n\right) = \sum_{i=1}^{n} \operatorname{var}\left(X_i\right) = \sum_{i=1}^{n} \sigma_X^2 = n\,\sigma_X^2$$

© M. Tummala & C. W. Therrien 2004

PDF of Sum of Two r.v.s

Let $Y = X_1 + X_2$ where

X_1 and X_2 are <u>*independent*</u> random variables.

The moment generating function of Y is

$$M_Y(s) = E\left[e^{sY}\right] = E\left[e^{s(X_i + X_2)}\right] = E\left[e^{sX_1} e^{sX_2}\right]$$

$$= E\left[e^{sX_1}\right] E\left[e^{sX_2}\right] \Leftarrow \text{because } X_1 \text{ and } X_2 \text{ are independent}$$

$$= M_{X_1}(s) M_{X_2}(s)$$

Therefore: $f_Y = f_{X_1} * f_{X_2}$

© M. Tummala & C. W. Therrien 2004

Generalization

Let $X_1 + X_2, \ldots, X_n$ be independent r.v.s

$$\boxed{Y = X_1 + X_2 + \cdots + X_n}$$

The moment generating function is

$$\begin{aligned} M_Y(s) &= E\left[e^{sY}\right] = E\left[e^{s(X_i + X_2 + \cdots + X_n)}\right] \\ &= E\left[e^{sX_1} e^{sX_2} \cdots e^{sX_n}\right] \\ &= E\left[e^{sX_1}\right] E\left[e^{sX_2}\right] \cdots E\left[e^{sX_n}\right] \\ &= M_{X_1}(s) M_{X_2}(s) \cdots M_{X_n}(s) \end{aligned}$$

Since $M_Y(s) \Leftrightarrow f_Y(y)$

$$\boxed{f_Y = f_{X_1} * f_{X_2} * \ldots * f_{X_n}}$$

© M. Tummala & C. W. Therrien 2004

<u>Example</u>:

Let $X_1 + X_2, \ldots, X_n$ be independent Gaussian random variables with m_i and σ_i^2. Let $Y = X_1 + X_2 + \ldots + X_n$, then

$$m_Y = m_1 + m_2 + \cdots + m_n \quad \text{and} \quad \sigma_Y^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$$

Find $f_Y(y)$.

The moment generating function converges only on the $j\omega$ axis. Therefore, let $s = j\omega$. The resulting characteristic function of X_i is

$$M_X(j\omega) = e^{-j\omega m_i + \omega^2 \sigma_i^2/2} \quad \Leftrightarrow \quad f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-(x_i - m_i)^2/2\sigma_i^2}$$

© M. Tummala & C. W. Therrien 2004

<u>Example (cont'd)</u>:

Then the characteristic function of Y is $Y = X_1 + X_2 + \ldots + X_n$

$$M_Y(j\omega) = M_{X_1}(j\omega) M_{X_2}(j\omega) \cdots M_{X_n}(j\omega)$$

$$= e^{-j\omega(m_1+m_2+\cdots+m_n)+\omega^2\left(\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2\right)/2}$$

By observation, we can write

$$f_Y(y) = \frac{1}{\sqrt{2\pi\left(\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2\right)}} e^{-\frac{(y-m_1-m_2-\cdots-m_n)^2}{2\left(\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2\right)}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma_Y} e^{-\frac{(y-m_y)^2}{2\sigma_Y^2}}$$

The sum of independent Gaussian random variables is Gaussian.

© M. Tummala & C. W. Therrien 2004

A New Random Variable (Discrete)

Recall: $f_K[k] = p(1-p)^k \quad k > 0$ (geometric)

Waiting time to first discrete event.

New: $f_K[k] * f_K[k] * \cdots * f_K[k]$ (m times)

$$f_K[k] = \binom{k-1}{m-1} p^m (1-p)^{k-m}$$ (Pascal or negative binomial)

Waiting time to m^{th} discrete event.

© M. Tummala & C. W. Therrien 2004

A New Random Variable (Continuous)

Recall: $f_X(x) = \lambda e^{-\lambda x}, \quad 0 \le x < \infty$ (exponential)

Describes waiting time to first event.

New: $f_X(x) * f_X(x) * \cdots * f_X(x)$ (m times)

$$f_X(x) = \frac{(\lambda x)^{m-1}}{(m-1)!} \lambda e^{-\lambda x}, \quad 0 \le x < \infty \quad (m\text{-Erlang})$$

Describes waiting time to m^{th} event.

© M. Tummala & C. W. Therrien 2004

<u>Example</u>: (4-Erlang)

Consider 4 components. One is used while the others are spares. Lifetime of each component is an exponential r.v. with an average lifetime of $1/\lambda = 2$ months. Find the probability that 4 components last more than 1 year.

Y_i = lifetime of a component

$X = Y_1 + Y_2 + Y_3 + Y_4$, combined lifetime of 4 components used in sequence

$$F_X(x) = \int_0^x \frac{(\lambda u)^{m-1}}{(m-1)!} \lambda e^{-\lambda u} du = 1 - \sum_{k=0}^{m-1} \frac{(\lambda x)^k}{k!} e^{-\lambda x}, \quad 0 \le x < \infty$$

$$\Pr[X > 12] = 1 - \Pr[X \le 12] = 1 - F(12)$$

$$= \sum_{k=0}^{3} \frac{(12/2)^k}{k!} e^{-12/2} = 0.1512$$

© M. Tummala & C. W. Therrien 2004

The Gamma Random Variable

pdf $$f_X(x) = \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} \qquad \begin{array}{c} 0 \le x < \infty \\ \alpha > 0, \quad \lambda > 0 \end{array}$$

where $$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx, \quad z > 0$$

$$\Gamma(m) = (m-1)! \qquad \text{if } m \text{ is a positive integer}$$

© M. Tummala & C. W. Therrien 2004

Special Cases of the Gamma RV

- Exponential. Let $\alpha = 1$

$$f_X(x) = \lambda e^{-\lambda x}, \quad 0 \le x < \infty$$

- m-Erlang. Let $\alpha = m$

$$f_X(x) = \frac{(\lambda x)^{m-1}}{(m-1)!} \lambda e^{-\lambda x}, \quad 0 \le x < \infty$$

- Chi-Square. Let $\lambda = 1/2$, $\alpha = k/2$

$$f_X(x) = \frac{\left(\frac{x}{2}\right)^{(k/2)-1}}{2\Gamma\left(\frac{k}{2}\right)} e^{-x/2}, \quad 0 \le x < \infty$$

© M. Tummala & C. W. Therrien 2004