Bi-dimensional Random variables

- Moments of a random variable X

$$E[x^n y^n] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^n y^n f_{X,Y}(x,y)dxdy$$

$$E[X] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f_{X,Y}(x,y)dxdy = \int_{-\infty}^{+\infty} x\left[\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\right]dydx = \int_{-\infty}^{+\infty} x f_X(x)dx$$

$$E[Y] = \int_{-\infty}^{+\infty} y f_y(y)dy$$

$E[XY] = E[X].E[Y]$ If X and Y are independents and in this case

$$f_{X,Y}(x,y) = f_X(x).f_Y(y)$$

Example: (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 $f_{X_2}(x_2)$

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

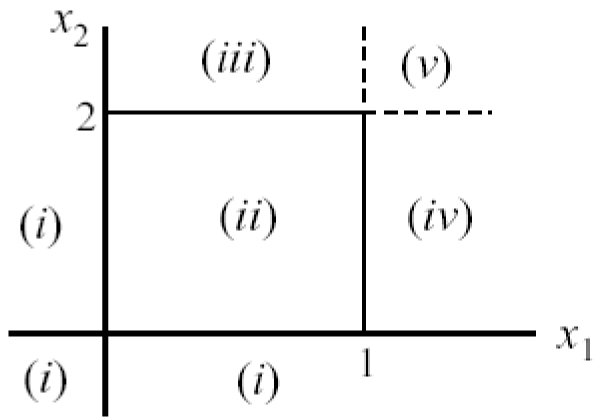
$$0\le x_2<\infty$$

Example:

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

Find the joint cdf.

$$F_{X_1,X_2}(x_1,x_2)=\Pr[X_1\le x_1, X_2\le x_2]=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1,X_2}(z_1,z_2)\,dz_2dz_1$$



Case i $\quad x_1<0$ or $x_2<0$ or both $x_1, x_2<0$

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

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

$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{x_1}\int_0^{x_2} dz_2 dz_1 = \frac{1}{2}x_1x_2$$

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

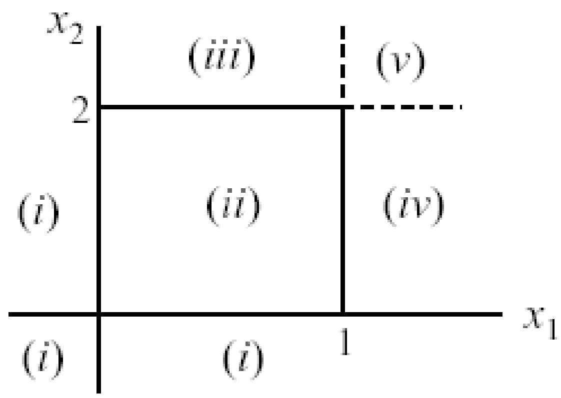
$$F_{X_1,X_2}(x_1,x_2) = \frac{1}{2}\int_0^{x_1}\int_0^{2} dz_2 dz_1 = x_1$$

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 $x_1 > 1, x_2 > 2$

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



Covariance

Definition: Let X and Y be jointly distributed with mean μ_X and μ_Y, respectively. The covariance of X and Y is defined as

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_Y)] \tag{426}$$

Theorem 5: Basic properties of the covariance

(a) $Cov(X,X) = Var(X) = \sigma_X^2$

(b) $Cov(X,Y) = Cov(Y,X)$

(c) $Cov(aX+b, cY+d) = ac\,Cov(X,Y)$

(d) $|Cov(X,Y)| \leq \sigma_X \sigma_Y$

(e) $Cov(X,Y) = E(XY) - E(X)E(Y)$

(f) If X and Y are independent, then $Cov(X,Y) = 0$

Covariance

Example 6: Express $Var(X+Y)$ in terms of σ_X^2, σ_Y^2 and $Cov(X,Y)$.

$$\begin{aligned} Var(X+Y) &= E[(X+Y)^2] - (E(X+Y))^2 \\ &= E(X^2+Y^2+2XY) - (\mu_X+\mu_Y)^2 \\ &= E(X^2)+E(Y^2)+2E(XY) \\ & \quad -\mu_X^2 - \mu_Y^2 - 2\mu_X\mu_Y \\ &= Var(X)+Var(Y)+2Cov(X,Y) \end{aligned}$$

Theorem 6:

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} Cov(X_i, X_j)$$

Correlation coefficient

Definition: The correlation coefficient of RVs X and Y is defined as

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

Remarks:

- The main difference between $\rho(X, Y)$ and $Cov(X, Y)$ is that the former has been normalized and is a dimensionless quantity.
- In fact, we have that $-1 \leq \rho(X, Y) \leq 1$

Example 8: Let RVs X and Y be jointly uniform over the region $D = \{(x, y) : 0 < x < y < 1\}$. Find $\rho(X, Y)$.

Correlation coefficient

Theorem 7: Basic properties of the correlation coefficient

(a) $\rho(X, X) = 1$

(b) $\rho(X, Y) = \rho(Y, X)$

(c) $\rho(aX + b, cY + d) = \epsilon\rho(X, Y)$ where $\epsilon = \text{sign}(ac)$.

(d) $|\rho(X, Y)| \leq 1$

(e) $\rho(X, Y) = 1 \Leftrightarrow Y = aX + b$ for some $a > 0$

$\rho(X, Y) = -1 \Leftrightarrow Y = aX + b$ for some $a < 0$

(f) If X and Y are independent, then $\rho(X, Y) = 0$

Discussion:

- The correlation coefficient $\rho(X, Y)$ provides a measure of the degree of linear association between RVs X and Y.
- If $\rho(X, Y) = 1$, then $Y = aX + b$ for some real numbers $a > 0$ and b. That is, let $L = \{(x, y) : y = ax + b\}$. Then $P((X, Y) \in L) = 1$:

Correlation coefficient

- If $0 < \rho(X,Y) < 1$, we have an intermediate situation: the joint pdf of X and Y is more or less concentrated along some line $L = \{(x,y) : y = ax + b\}$ with positive slope $a > 0$.

- Standard terminology:
  - if $\rho(X,Y) > 0$, we say that X and Y are positively correlated
  - if $\rho(X,Y) < 0$, we say that X and Y are negatively correlated
  - if $\rho(X,Y) = 0$, we say that X and Y are uncorrelated

Second moment (**correlation matrix**)

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

$$= E\begin{bmatrix} X_1 X_1 & X_1 X_2 \\ X_2 X_1 & X_2 X_2 \end{bmatrix}$$

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

When the correlation of X_1 and X_2 is defined as

$$r_{ij} = E\left[X_i X_j\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i x_j f_{X_1 X_2}\left(x_1, x_2\right) dx_1 dx_2, \qquad i = 1, 2;\ j = 1, 2$$

Second central moment (**covariance matrix**)

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

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

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

where the covariance of X_1 and X_2 is defined as

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

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

An important relation:

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

Proof:

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

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

Bivariate Gaussian Random Variables X_1, X_2

Let X_1 and X_2 be jointly Gaussian

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

Gaussian pdf for an $N \times 1$ vector

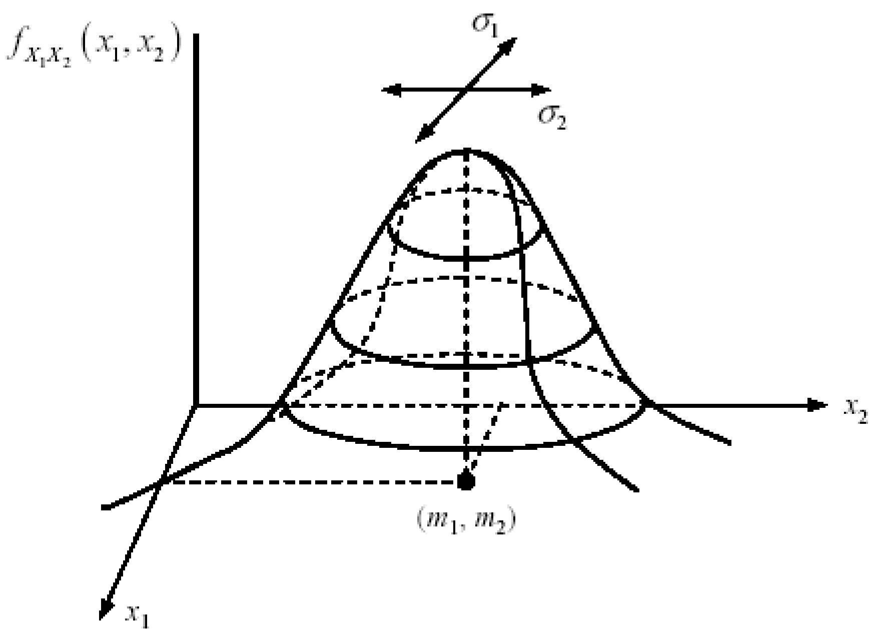
$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{(2\pi)^{N/2} \left| C_X \right|^{1/2}} e^{-1/2(\mathbf{x}-\mathbf{m}_X)^T C_X^{-1} (\mathbf{x}-\mathbf{m}_X)}$$

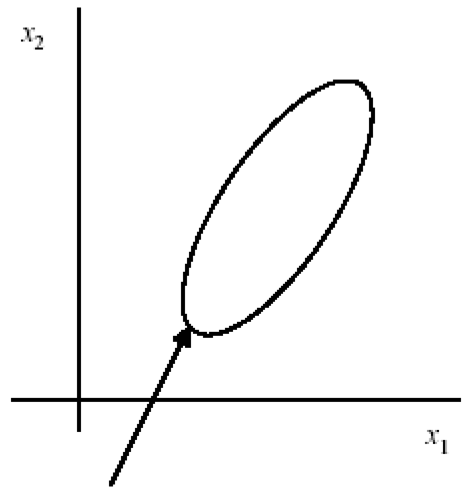
For $N = 2$ and by letting $\sigma_1^2 = \sigma_2^2 = \sigma^2$,

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

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

Bivariate Gaussian PDF (Joint Gaussian density function)





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

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

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 $\mathbf{m}_X = 0$, $\sigma_1^2 = \sigma_2^2 = \sigma^2$

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

Expectation of Two Random Variables (Random Vectors)

Let us use the compact notation

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

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

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

Mean of X **(mean vector)**

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