Financial Engineering & Risk Management

Review of Multivariate Distributions

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Multivariate Distributions I

Let $\mathbf{X} = (X_1 \ \ldots \ X_n)^\top$ be an n-dimensional vector of random variables.

Definition. For all $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the joint cumulative distribution function (CDF) of $\mathbf{X}$ satisfies

$$F_{\mathbf{X}}(\mathbf{x}) \ = \ F_{\mathbf{X}}(x_1, \ldots, x_n) \ = \ \mathsf{P}(X_1 \le x_1, \ldots, X_n \le x_n).$$

Definition. For a fixed i, the marginal CDF of X_i satisfies

$$F_{X_i}(x_i) \ = \ F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots \infty).$$

It is straightforward to generalize the previous definition to joint marginal distributions. For example, the joint marginal distribution of X_i and X_j satisfies

$$F_{ij}(x_i, x_j) \ = \ F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty, x_j, \infty, \ldots \infty).$$

We also say that $\mathbf{X}$ has joint PDF $f_{\mathbf{X}}(\cdot, \ldots, \cdot)$ if

$$F_{\mathbf{X}}(x_1, \ldots, x_n) \ = \ \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(u_1, \ldots, u_n) \, du_1 \ldots du_n.$$

Multivariate Distributions II

Definition. If $\mathbf{X_1} = (X_1, \ldots X_k)^\top$ and $\mathbf{X_2} = (X_{k+1} \ldots X_n)^\top$ is a partition of $\mathbf{X}$ then the conditional CDF of $\mathbf{X_2}$ given $\mathbf{X_1}$ satisfies

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2} \,|\, \mathbf{x_1}) \;=\; P(\mathbf{X_2} \leq \mathbf{x_2} \,|\, \mathbf{X_1} = \mathbf{x_1}).$$

If $\mathbf{X}$ has a PDF, $f_{\mathbf{X}}(\cdot)$, then the conditional PDF of $\mathbf{X_2}$ given $\mathbf{X_1}$ satisfies

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2 \,|\, \mathbf{x}_1) \;=\; \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; \frac{f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1 \,|\, \mathbf{x}_2) f_{\mathbf{X}_2}(\mathbf{x}_2)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \tag{1}$$

and the conditional CDF is then given by

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2}|\mathbf{x_1}) \;=\; \int_{-\infty}^{x_{k+1}} \cdots \int_{-\infty}^{x_n} \frac{f_{\mathbf{X}}(x_1, \ldots, x_k, u_{k+1}, \ldots, u_n)}{f_{\mathbf{X_1}}(\mathbf{x_1})} \; du_{k+1} \ldots du_n$$

where $f_{\mathbf{X_1}}(\cdot)$ is the joint marginal PDF of $\mathbf{X_1}$ which is given by

$$f_{\mathbf{X_1}}(x_1, \ldots, x_k) \;=\; \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, \ldots, x_k, u_{k+1}, \ldots, u_n) \; du_{k+1} \ldots du_n.$$

Independence

Definition. We say the collection $\mathbf{X}$ is independent if the joint CDF can be factored into the product of the marginal CDFs so that

$$F_{\mathbf{X}}(x_1, \ldots, x_n) \;=\; F_{X_1}(x_1) \ldots F_{X_n}(x_n).$$

If $\mathbf{X}$ has a PDF, $f_{\mathbf{X}}(\cdot)$ then independence implies that the PDF also factorizes into the product of marginal PDFs so that

$$f_{\mathbf{X}}(\mathbf{x}) \;=\; f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

Can also see from (1) that if $\mathbf{X_1}$ and $\mathbf{X_2}$ are independent then

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2 \,|\, \mathbf{x}_1) \;=\; \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; \frac{f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; f_{\mathbf{X}_2}(\mathbf{x}_2)$$

– so having information about $\mathbf{X}_1$ tells you nothing about $\mathbf{X}_2$.

Implications of Independence

Let X and Y be independent random variables. Then for any events, A and B,

$$\mathsf{P}(X \in A,\ Y \in B) = \mathsf{P}(X \in A)\,\mathsf{P}(Y \in B) \tag{2}$$

More generally, for any function, $f(\cdot)$ and $g(\cdot)$, independence of X and Y implies

$$\mathsf{E}[f(X)g(Y)] = \mathsf{E}[f(X)]\mathsf{E}[g(Y)]. \tag{3}$$

In fact, (2) follows from (3) since

$$\begin{aligned} \mathsf{P}(X \in A,\ Y \in B) &= \mathsf{E}\left[1_{\{X\in A\}}1_{\{Y\in B\}}\right] \\ &= \mathsf{E}\left[1_{\{X\in A\}}\right]\mathsf{E}\left[1_{\{Y\in B\}}\right] \qquad \text{by } (3) \\ &= \mathsf{P}(X \in A)\,\mathsf{P}(Y \in B). \end{aligned}$$

Implications of Independence

More generally, if $X_1, \ldots X_n$ are independent random variables then

$$\mathsf{E}\left[f_1(X_1)f_2(X_2)\cdots f_n(X_n)\right] \;=\; \mathsf{E}[f_1(X_1)]\mathsf{E}[f_2(X_2)]\cdots\mathsf{E}[f_n(X_n)].$$

Random variables can also be conditionally independent. For example, we say X and Y are conditionally independent given Z if

$$\mathsf{E}[f(X)g(Y)\,|\,Z] \;=\; \mathsf{E}[f(X)\,|\,Z]\;\mathsf{E}[g(Y)\,|\,Z].$$

– used in the (in)famous Gaussian copula model for pricing CDOs!

In particular, let D_i be the event that the i^{th} bond in a portfolio defaults.

Not reasonable to assume that the D_i's are independent. Why?

But maybe they are conditionally independent given Z so that

$$\mathsf{P}(D_1, \ldots, D_n\,|\,Z) \;=\; \mathsf{P}(D_1\,|\,Z)\cdots\mathsf{P}(D_n\,|\,Z)$$

– often easy to compute this.

The Mean Vector and Covariance Matrix

The mean vector of $\mathbf{X}$ is given by

$$\mathsf{E}[\mathbf{X}] \;:=\; \left(\mathsf{E}[X_1] \;\ldots\; \mathsf{E}[X_n]\right)^\top$$

and the covariance matrix of $\mathbf{X}$ satisfies

$$\mathbf{\Sigma} := \mathsf{Cov}(\mathbf{X}) \;:=\; \mathsf{E}\left[(\mathbf{X} - \mathsf{E}[\mathbf{X}])\,(\mathbf{X} - \mathsf{E}[\mathbf{X}])^\top\right]$$

so that the $(i,j)^{th}$ element of $\mathbf{\Sigma}$ is simply the covariance of X_i and X_j.

The covariance matrix is symmetric and its diagonal elements satisfy $\mathbf{\Sigma}_{i,i} \geq 0$.

It is also positive semi-definite so that $\mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

The correlation matrix, $\rho(\mathbf{X})$, has $(i,j)^{th}$ element $\rho_{ij} := \mathsf{Corr}(X_i, X_j)$

- it is also symmetric, positive semi-definite and has 1's along the diagonal.

Variances and Covariances

For any matrix $\mathbf{A} \in \mathbb{R}^{k \times n}$ and vector $\mathbf{a} \in \mathbb{R}^k$ we have

$$\begin{aligned} \mathsf{E}\left[\mathbf{AX} + \mathbf{a}\right] &= \mathbf{A}\mathsf{E}\left[\mathbf{X}\right] + \mathbf{a} && (4) \\ \mathsf{Cov}(\mathbf{AX} + \mathbf{a}) &= \mathbf{A}\,\mathsf{Cov}(\mathbf{X})\,\mathbf{A}^\top. && (5) \end{aligned}$$

Note that (5) implies

$$\mathsf{Var}(aX + bY) = a^2\mathsf{Var}(X) + b^2\mathsf{Var}(Y) + 2ab\,\mathsf{Cov}(X, Y).$$

If X and Y independent, then $\mathsf{Cov}(X, Y) = 0$

– but converse not true in general.

Financial Engineering & Risk Management

The Multivariate Normal Distribution

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

The Multivariate Normal Distribution I

If the n-dimensional vector $\mathbf{X}$ is multivariate normal with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ then we write

$$\mathbf{X} \sim \mathsf{MN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

The PDF of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) \;=\; \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}}\; e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top\, \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

where $|\cdot|$ denotes the determinant.

Standard multivariate normal has $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I_n}$, the $n \times n$ identity matrix

- in this case the X_i's are independent.

The moment generating function (MGF) of $\mathbf{X}$ satisfies

$$\phi_{\mathbf{X}}(\mathbf{s}) \;=\; \mathsf{E}\left[e^{\mathbf{s}^\top\mathbf{X}}\right] \;=\; e^{\mathbf{s}^\top\boldsymbol{\mu}+\frac{1}{2}\mathbf{s}^\top\boldsymbol{\Sigma}s}.$$

The Multivariate Normal Distribution II

Recall our partition of $\mathbf{X}$ into $\mathbf{X_1} = (X_1 \ \ldots \ X_k)^\top$ and $\mathbf{X_2} = (X_{k+1} \ \ldots \ X_n)^\top$. Can extend this notation naturally so that

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}.$$

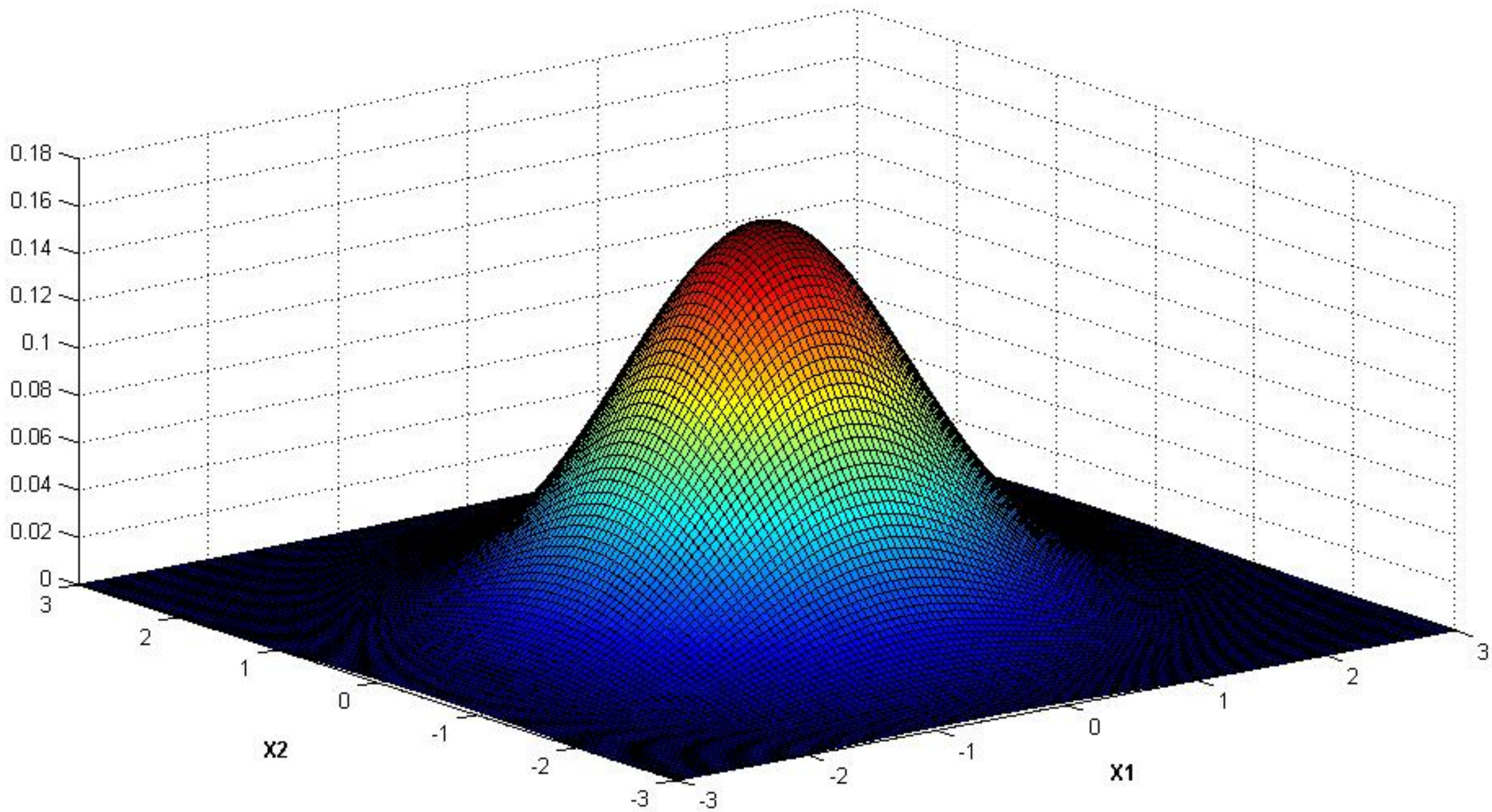
are the mean vector and covariance matrix of $(\mathbf{X_1}, \mathbf{X_2})$.

Then have following results on marginal and conditional distributions of $\mathbf{X}$:

Marginal Distribution

The marginal distribution of a multivariate normal random vector is itself normal. In particular, $\mathbf{X_i} \sim \mathrm{MN}(\boldsymbol{\mu}_i, \mathbf{\Sigma}_{ii})$, for $i = 1, 2$.

The Bivariate Normal PDF



The Bivariate Normal PDF

The Multivariate Normal Distribution III

Conditional Distribution

Assuming $\mathbf{\Sigma}$ is positive definite, the conditional distribution of a multivariate normal distribution is also a multivariate normal distribution. In particular,

$$\mathbf{X_2} \mid \mathbf{X_1} = \mathbf{x_1} \;\sim\; \mathsf{MN}(\boldsymbol{\mu}_{2.1}, \mathbf{\Sigma}_{2.1})$$

where $\boldsymbol{\mu}_{2.1} = \mu_2 + \mathbf{\Sigma}_{21}\, \mathbf{\Sigma}_{11}^{-1}\, (\mathbf{x_1} - \boldsymbol{\mu}_1)$ and $\mathbf{\Sigma}_{2.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$.

Linear Combinations

A linear combination, $\mathbf{AX} + \mathbf{a}$, of a multivariate normal random vector, $\mathbf{X}$, is normally distributed with mean vector, $\mathbf{A}\mathsf{E}\left[\mathbf{X}\right] + \mathbf{a}$, and covariance matrix, $\mathbf{A}\ \mathsf{Cov}(\mathbf{X})\ \mathbf{A}^\top$.

Financial Engineering & Risk Management

Introduction to Martingales

M. Haugh **G. Iyengar**

Department of Industrial Engineering and Operations Research
Columbia University

Martingales

Definition. A random process, $\{X_n \; : \; 0 \leq n \leq \infty\}$, is a martingale with respect to the information filtration, $\mathcal{F}_n$, and probability distribution, P, if

1. $\mathsf{E}^P[|X_n|] < \infty$ for all $n \geq 0$
2. $\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] = X_n$ for all $n, m \geq 0$.

Martingales are used to model fair games and have a rich history in the modeling of gambling problems.

We define a submartingale by replacing condition #2 with

$$\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] \geq X_n \quad \text{for all } n, m \geq 0.$$

And we define a supermartingale by replacing condition #2 with

$$\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] \leq X_n \quad \text{for all } n, m \geq 0.$$

A martingale is both a submartingale and a supermartingale.

Constructing a Martingale from a Random Walk

Let $S_n := \sum_{i=1}^n X_i$ be a random walk where the X_i's are IID with mean μ.

Let $M_n := S_n - n\mu$. Then M_n is a martingale because:

$$\begin{aligned}
\mathsf{E}_n[M_{n+m}] &= \mathsf{E}_n\left[\sum_{i=1}^{n+m} X_i - (n+m)\mu\right] \\
&= \mathsf{E}_n\left[\sum_{i=1}^{n+m} X_i\right] - (n+m)\mu \\
&= \sum_{i=1}^{n} X_i \; + \; \mathsf{E}_n\left[\sum_{i=n+1}^{n+m} X_i\right] - (n+m)\mu \\
&= \sum_{i=1}^{n} X_i \; + \; m\mu - (n+m)\mu \; = \; M_n.
\end{aligned}$$

A Martingale Betting Strategy

Let $X_1, X_2, \ldots$ be IID random variables with

$$\mathsf{P}(X_i = 1) \;=\; \mathsf{P}(X_i = -1) \;=\; \frac{1}{2}.$$

Can imagine X_i representing the result of coin-flipping game:

- Win $\$1$ if coin comes up heads
- Lose $\$1$ if coin comes up tails

Consider now a doubling strategy where we keep doubling the bet until we eventually win. Once we win, we stop and our initial bet is $\$1$.

First note that size of bet on n^{th} play is 2^{n-1}

– assuming we're still playing at time n.

Let W_n denote total winnings after n coin tosses assuming $W_0 = 0$.

Then W_n is a martingale!

A Martingale Betting Strategy

To see this, first note that $W_n \in \{1, -2^n + 1\}$ for all n. Why?

1. Suppose we win for first time on n^{th} bet. Then

$$\begin{aligned} W_n &= -\left(1 + 2 + \cdots + 2^{n-2}\right) \; + \; 2^{n-1} \\ &= -\left(2^{n-1} - 1\right) \; + \; 2^{n-1} \\ &= 1 \end{aligned}$$

2. If we have not yet won after n bets then

$$\begin{aligned} W_n &= -\left(1 + 2 + \cdots + 2^{n-1}\right) \\ &= -2^n + 1. \end{aligned}$$

To show W_n is a martingale only need to show $\mathrm{E}[\, W_{n+1} \mid W_n] = W_n$

– then follows by iterated expectations that $\mathrm{E}[\, W_{n+m} \mid W_n] = W_n$.

A Martingale Betting Strategy

There are two cases to consider:

1: $W_n = 1$: then $\mathsf{P}(W_{n+1} = 1 | W_n = 1) = 1$ so

$$\mathsf{E}[W_{n+1} \mid W_n = 1] = 1 = W_n \tag{6}$$

2: $W_n = -2^n + 1$: bet 2^n on $(n+1)^{th}$ toss so $W_{n+1} \in \{1, -2^{n+1} + 1\}$.
Clear that

$$\begin{aligned} \mathsf{P}(W_{n+1} = 1 \mid W_n = -2^n + 1) &= 1/2 \\ \mathsf{P}(W_{n+1} = -2^{n+1} + 1 \mid W_n = -2^n + 1) &= 1/2 \end{aligned}$$

so that

$$\begin{aligned} \mathsf{E}[W_{n+1} \mid W_n = -2^n + 1] &= (1/2)1 + (1/2)(-2^{n+1} + 1) \\ &= -2^n + 1 = W_n. \end{aligned} \tag{7}$$

From (6) and (7) we see that $\mathsf{E}[W_{n+1} \mid W_n] = W_n$.

Polya's Urn

Consider an urn which contains red balls and green balls.
Initially there is just one green ball and one red ball in the urn.

At each time step a ball is chosen randomly from the urn:

1. If ball is red, then it's returned to the urn with an additional red ball.
2. If ball is green, then it's returned to the urn with an additional green ball.

Let X_n denote the number of red balls in the urn after n draws. Then

$$\begin{aligned} \mathsf{P}(X_{n+1} = k+1 \,|\, X_n = k) &= \frac{k}{n+2} \\ \mathsf{P}(X_{n+1} = k \,|\, X_n = k) &= \frac{n+2-k}{n+2}. \end{aligned}$$

Show that $M_n := X_n/(n+2)$ is a martingale.

(These martingale examples taken from *"Introduction to Stochastic Processes"* (Chapman & Hall) by Gregory F. Lawler.)