The Univariate Normal Distribution

- $X \sim \mathsf{N}(\mu, \sigma^2)$, check Appendix D for its pdf and mgf,

$$\mathbb{E}(X) = \mu, \quad \mathrm{Var}(X) = \sigma^2.$$

  How to compute the mgf $M_X(t)$?

- $Z \sim \mathsf{N}(0,1)$ is called the standard normal rv. $\Phi(z)$ denotes its CDF.

$$\Phi(-z) = 1 - \Phi(z), \quad z > 0.$$

- Use the MGF approach, we can show that: if $X \sim \mathsf{N}(\mu, \sigma^2)$, then $aX + b \sim \mathsf{N}(a\mu + b, a^2\sigma^2)$. Specially, $\frac{1}{\sigma}(X - \mu) \sim N(0,1)$.

- Use the MGF approach, we can show that: Linear combinations of independent normal rv's are normal (Thm 3.4.2): Let $X_1, \ldots, X_n$ be independent random variables with $X_i \sim \mathsf{N}(\mu_i, \sigma_i^2)$, then

$$\sum_{i=1}^n a_i X_i \sim \mathsf{N}\Big(\sum_i a_i \mu_i, \sum_i a_i^2 \sigma_i^2\Big).$$

- Distributions related to Normal

  - Z_i iid $\sim N(0,1)$, then $Z_1^2 + \cdots + Z_n^2 \sim \chi_n^2$.

$$W \sim \chi_n^2, \qquad \mathbb{E}(W) = n, \qquad \mathrm{Var}(W) = 2n.$$

  - $Z \sim N(0,1)$ and $W \sim \chi_n^2$ are independent, then

$$\frac{Z}{\sqrt{W/n}} \sim t_n, ,$$

    the student t-dist with df n (df = degree of freedom). The pdf for t_n looks very much like a standard normal curve: bell-shaped and centered at zero, however, the t curve has a fatter tail than the normal and t_n approaches to $\mathsf{N}(0,1)$ when n goes to infinity. Check the wiki page.

  - Check the examples in `[BivariateNormalDistribution2]`.

Bivariate Normal Distribution

(X_1, X_2) follows a bivariate normal distribution and their joint pdf is given by

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\Big(\frac{x_1-\mu_1}{\sigma_1}\Big)^2 - 2\rho\Big(\frac{x_1-\mu_1}{\sigma_1}\Big)\Big(\frac{x_2-\mu_2}{\sigma_2}\Big) + \Big(\frac{x_2-\mu_2}{\sigma_2}\Big)^2 \right] \right\}$$

with $-\infty < x_1, x_2 < \infty$, where μ_1 and σ_1^2 are the mean and variance, respectively, of X_1; μ_2 and σ_2^2 are the mean and variance, respectively, of X_2; $-1 < \rho < 1$ is the correlation coefficient[1].

Let

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where we have assumed that $\rho^2 < 1$, then we have

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix},$$

and the joint pdf $f(x_1, x_2)$ can also be written as

$$\frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}.$$

- Marginals: $X_i \sim \mathsf{N}(\mu_i, \sigma_i^2)$ where $i = 1, 2$.
- Conditionals:

$$X_1 \mid X_2 = x_2 \sim \mathsf{N}\Big(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), (1-\rho^2)\sigma_1^2\Big)$$

$$X_2 \mid X_1 = x_1 \sim \mathsf{N}\Big(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), (1-\rho^2)\sigma_2^2\Big)$$

- Linear Combinations:

$$aX_1 + bX_2 \sim \mathsf{N}\big(a\mu_1 + b\mu_2, a^2\sigma_1^2 + 2ab\rho\sigma_1\sigma_2 + b^2\sigma_2^2\big)$$

- Uncorrelated = Independent: $\rho = 0$ implies X_1 and X_2 are independent.
- Note: All the statements above assume that the joint distribution of (X_1, X_2) is normal. However,

$$X_1 \sim \mathsf{Norm}, \quad X_2 \sim \mathsf{Norm} \quad \text{does NOT imply} \quad (X_1, X_2) \sim \mathsf{Norm}.$$

  So if X_1 and X_2 are marginally normally distributed with correlation 0, we cannot conclude that X_1 and X_2 are independent.

  Example: Suppose $X \sim \mathsf{N}(0, 1)$. Then generate the other random variable Y by tossing a fair coin: let $Y = X$ if we observe a head, otherwise $Y = -X$. In other words,

$$Y = WX, \quad \text{where } W = \begin{cases} 1 & \text{with prob } 1/2, \\ -1 & \text{with prob } 1/2. \end{cases}$$

  - You can show that 1) $Y \sim \mathsf{N}(0, 1)$, and 2) X and Y are uncorrelated.
  - You can also show that X and Y are not independent, which, consequently, implies that the joint pdf of (X, Y) is not normal. *Hint:* check whether $P(Y > 2 \mid X = 2) = P(Y > 2)$.

[1]That is, $\sigma_{12} = \rho\sigma_1\sigma_2$ is the covariance of X_1 and X_2.

The Multivariate Normal Distribution

- Let $\mathbf{Z} = (Z_1, \ldots, Z_n)$ where Z_i's are iid $\sim \mathsf{N}(0,1)$ rv's. Then $\mathbf{Z}$ follows a multivariate normal distribution, denoted by $\mathsf{N}_n(\mathbf{0}, \mathbf{I}_n)$, with

$$f(\mathbf{z}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-z_i^2} = \frac{1}{(2\pi)^{n/2}} \exp\Big\{-\frac{1}{2}\sum_{i=1}^n z_i^2\Big\} = \frac{1}{(2\pi)^{n/2}} \exp\Big\{-\frac{1}{2}\mathbf{z}^t\mathbf{z}\Big\},$$

$$M(\mathbf{t}) = \mathbb{E}[\exp\{\mathbf{t}^t\mathbf{Z}\}] = \exp\Big\{\frac{1}{2}\mathbf{t}^t\mathbf{t}\Big\},$$

  and

$$\mathbb{E}(\mathbf{Z}) = \mathbf{0}, \qquad \mathrm{Cov}(\mathbf{Z}) = \mathbf{I}_n.$$

- (Def 3.5.1) X has a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance Σ, denoted by $\mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$ if

$$M_X(\mathbf{t}) = \exp\Big\{\mathbf{t}^t\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^t\Sigma\mathbf{t}\Big\}.$$

  Why don't we define X via its pdf?

- Recall the definition of the covariance matrix for a random vector in Sec 2.6.1. Any covariance matrix Σ should be symmetric and *positive semi-definite* (psd), where psd means

$$\mathbf{a}^t\Sigma\mathbf{a} \geq 0.$$

  Any symmetric psd matrix has a *spectral decomposition*

$$\Sigma = \Gamma^t\Lambda\Gamma, \qquad \Lambda = \mathrm{diag}(\lambda_i)_{i=1}^n,$$

  and $\Gamma_{n\times n}$ is a orthonormal matrix, i.e., $\Gamma\Gamma^t = \mathbf{I}_n$.

- If $\lambda_n > 0$, then Σ is of full rank, so $|\Sigma| > 0$ and Σ^{-1} exists. The pdf of $\mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$ is given by

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}.$$

- Properties of Multivariate Normal rv's
  - (Thm 3.5.1) Affine transformations of a normal are still normal:

$$Y \sim \mathsf{N}_n(\boldsymbol{\mu}, \Sigma) \Longrightarrow A_{m\times n}Y + b_{m\times 1} \sim \mathsf{N}_m(A\boldsymbol{\mu} + b, A\Sigma A^t).$$

  - (Corollary 3.5.1) Marginals of a normal are still normal.
  - (Thm 3.5.3) Conditionals of a normal are still normal.

$$\mathbf{X}_1|\mathbf{X}_2 \sim \mathsf{N}_m\Big(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Big)$$

  - (Thm 3.5.2) For multivariate normals, uncorrelated = independent.

Example 2: $\mathbf{X} = (X_1, X_2, X_3)$ follows a multivariate normal distribution with mean $(5, 3, 7)^t$ and covariance matrix

$$\Sigma = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{pmatrix}.$$

- Find $\mathbb{P}(X_1 > 8)$. ($1 - \Phi(1.5) = 0.0668$)
- Find $\mathbb{P}(X_1 > 8 | X_2 = 1, X_3 = 10)$. $X_1 | X_2 = 1, X_3 = 10 \sim \mathsf{N}(5.75, 3.7188)$
- Find $\mathbb{P}(4X_1 - 3X_2 + 5X_3 < 63)$. $4X_1 - 3X_2 + 5X_3 \sim \mathsf{N}(46, 289)$

$Y = 4X_1 - 3X_2 + 5X_3$ follows a normal distribution. How to find the corresponding mean and variance?

$$\mu_Y = (4, -3, 5) \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = 4\mu_1 - 3\mu_2 + 5\mu_3 = 4(5) - 3(3) + 5(7) = 46.$$

$$\sigma_Y^2 = (4, -3, 5) \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{pmatrix} \begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix},$$

OR you can calculate σ_Y^2 this way: recall that $Y = 4X_1 - 3X_2 + 5X_3$ is a linear combination of X_1, X_2, X_3. I'll call the three numbers, 4, -3, and 5 weights. Put the three weights $(4, -3, 5)$ on top of the matrix Σ (as an additional row) and also along the side of Σ (as the 1st column) as the following

	4	-3	5
4	4	-1	0
-3	-1	4	2
5	0	2	9

Then sum over $3 \times 3 = 9$ terms and each term is one entry in the covariance matrix (in black) times the corresponding column weight (in blue) and row weight (in blue), i.e.,

$$\begin{aligned} \sigma_Y^2 &= 4(4)(4) + (-1)(-3)(4) + 0(5)(4) \\ &\quad +(-1)(4)(-3) + 4(-3)(-3) + 2(5)(-3) \\ &\quad +0(4)(5) + 2(-3)(5) + 9(5)(5) = 289. \end{aligned}$$

- Find ρ_{12}, ρ_{13} and ρ_{23}.