STAT 4100 Lecture Note

Week Two (Sep 12, 14 & 16, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Sep/09 10:42:46

Multivariate distribution

- Random vector $\mathbf{X} = (X_1, \ldots, X_n)$ with realization $\boldsymbol{x} = (x_1, \ldots, x_n)$
- Discrete
  - Joint pmf
$$p_{\mathbf{X}}(\boldsymbol{x}) = \Pr(X_1 = x_1, \ldots, X_n = x_n)$$
  - $\operatorname{supp}(\mathbf{X}) = \operatorname{supp}(p_{\mathbf{X}}) = \{\boldsymbol{x} \in \mathbb{R}^n : p_{\mathbf{X}}(\boldsymbol{x}) > 0\}$
  - Marginal pmf of $(X_1, \ldots, X_k)$
$$p_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \sum_{(x_{k+1},\ldots,x_n)\in\mathbb{R}^{n-k}} p_{\mathbf{X}}(\boldsymbol{x})$$

- Continuous
  - Joint pdf $f_{\mathbf{X}}(\boldsymbol{x})$ such that $\Pr(\mathbf{X} \in B) = \int_B f_{\mathbf{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ for each Borel set $B \subset \mathbb{R}^n$
  - $\operatorname{supp}(\mathbf{X}) = \operatorname{supp}(f_{\mathbf{X}}) = \{\boldsymbol{x} \in \mathbb{R}^n : f_{\mathbf{X}}(\boldsymbol{x}) > 0\}$
  - Marginal pdf of $(X_1, \ldots, X_k)$
    * $f_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \int_{\mathbb{R}^{n-k}} f_{\mathbf{X}}(\boldsymbol{x})\mathrm{d}x_{k+1}\cdots\mathrm{d}x_n$

Find the joint pdf of random vector $\mathbf{Y} = \boldsymbol{g}(\mathbf{X})$ by multivariate transformation (CB Sec. 4.3 & 4.6)

- Conditions
  - $\mathbf{X}$ and $\mathbf{Y}$ both of n dimensions
  - $\boldsymbol{g}(\cdot) = (g_1(\cdot), \ldots, g_n(\cdot)) : \operatorname{supp}(\mathbf{X}) \to \operatorname{supp}(\mathbf{Y})$ is one-to-one, i.e.,
    * $\boldsymbol{y} = (y_1, \ldots, y_n) = (g_1(\boldsymbol{x}), \ldots, g_n(\boldsymbol{x})) = \boldsymbol{g}(\boldsymbol{x})$
    * $\boldsymbol{x} = (x_1, \ldots, x_n) = \boldsymbol{g}^{-1}(\boldsymbol{y}) = (g_1^{-1}(\boldsymbol{y}), \ldots, g_n^{-1}(\boldsymbol{y}))$
- Jacobian matrices
  - Jacobian matrix of transformation $\boldsymbol{g}^{-1}$
$$\mathbf{J}_{\boldsymbol{g}^{-1}} = \mathbf{J}_{\boldsymbol{g}^{-1}}(\boldsymbol{y}) = \left[\frac{\partial g_i^{-1}(\boldsymbol{y})}{\partial y_j}\right]_{n\times n} = \begin{bmatrix} \frac{\partial g_1^{-1}(\boldsymbol{y})}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}(\boldsymbol{y})}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}(\boldsymbol{y})}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}(\boldsymbol{y})}{\partial y_n} \end{bmatrix}$$
  - Jacobian matrix of transformation $\boldsymbol{g}$
$$\mathbf{J}_{\boldsymbol{g}} = \mathbf{J}_{\boldsymbol{g}}(\boldsymbol{x}) = \left[\frac{\partial g_i(\boldsymbol{x})}{\partial x_j}\right]_{n\times n} = \begin{bmatrix} \frac{\partial g_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\boldsymbol{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_n(\boldsymbol{x})}{\partial x_n} \end{bmatrix}$$

  - $\mathbf{J}_{\boldsymbol{g}^{-1}}(\boldsymbol{y}) = \{\mathbf{J}_{\boldsymbol{g}}(\boldsymbol{g}^{-1}(\boldsymbol{y}))\}^{-1}$
    - Alternative way to reach $\mathbf{J}_{\boldsymbol{g}^{-1}}(\boldsymbol{y})$
- Then
$$f_{\mathbf{Y}}(\boldsymbol{y}) = f_{\mathbf{X}}\{g^{-1}(\boldsymbol{y})\}|\det\{\mathbf{J}_{\boldsymbol{g}^{-1}}(\boldsymbol{y})\}|\mathbf{1}_{\mathrm{supp}(\mathbf{Y})}(\boldsymbol{y}).$$
  - Never miss $\mathbf{1}_{\mathrm{supp}(\mathbf{Y})}(\boldsymbol{y})$
- If $\boldsymbol{g}$ is NOT one-to-one, one may figure out the cdf of $\mathbf{Y}$ and then differentiate it.

Example Lec3.1

X_1 and X_2 are iid from $\mathcal{N}(0,1)$. Find the joint pdf of $Y_1 = (X_1 + X_2)/\sqrt{2}$ and $Y_2 = (X_1 - X_2)/\sqrt{2}$ and show their independence.

Note: the sample mean and standard deviation are respectively $\bar{X} = (X_1 + X_2)/2 = Y_1/\sqrt{2}$ and $S = \sqrt{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2} = |Y_2|$.

Find the marginal pdf

1. Figure out the joint pdf first
2. Taking the Integral

Example Lec3.2

X_1 and X_2 are iid from $\mathcal{N}(0,1)$. Find the pdf of $U = \sqrt{X_1^2 + X_2^2}$.

Basics on square matrices

- Eigen-decomposition
  - $\mathbf{A}$ is a real $n \times n$ matrix
  - Eigenvalues of $\mathbf{A}$, say $\lambda_1 \geq \cdots \geq \lambda_n$: n roots of characteristic equation $\det(\lambda\mathbf{I}_n - \mathbf{A}) = 0$
  - The ith (Right) eigenvector $\boldsymbol{v}_i$: $\mathbf{A}\boldsymbol{v}_i = \lambda_i\boldsymbol{v}_i$
  - Eigen-decomposition: $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1}$
    - $\mathbf{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ both $n \times n$ matrices
    - Specifically $\mathbf{V}^{-1} = \mathbf{V}^\top$ for symmetric $\mathbf{A}$
  - Numerical implementation in *R*: eigen()
- Square root of matrices
  - $\mathbf{A}^{1/2} = \mathbf{V}\Lambda^{1/2}\mathbf{V}^\top$ if for semi-positive definite $\mathbf{A}$
    - Semi-positive/non-negative definite: symmetric $\mathbf{A}$ with eigenvalues all non-negative
    - $\Lambda^{1/2} = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$
    - $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$
  - $\mathbf{A}^{-1/2} = \mathbf{V}\Lambda^{-1/2}\mathbf{V}^\top$ for positive definite $\mathbf{A}$
    - Positive definite: symmetric $\mathbf{A}$ with eigenvalues all positive
    - $\Lambda^{-1/2} = \mathrm{diag}(\lambda_1^{-1/2}, \ldots, \lambda_n^{-1/2})$
    - $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$ and $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{I}$

- Determinant
  - $\det\mathbf{A} = \prod_{i=1}^n \lambda_i$
  - $\det(\mathbf{A}^\top) = \det\mathbf{A}$
  - $\det(\mathbf{A}^{-1}) = (\det\mathbf{A})^{-1}$
  - $\det(c\mathbf{A}) = c^n\det\mathbf{A}$ for $n \times n$ matrix $\mathbf{A}$ and scalar c
  - $\det(\mathbf{AB}) = (\det\mathbf{A})(\det\mathbf{B})$ for squared $\mathbf{A}$ and $\mathbf{B}$
- Trace
  - $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$
  - $\mathrm{tr}(c\mathbf{A}) = c\,\mathrm{tr}(\mathbf{A})$ for scalar c

- $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ for squared $\mathbf{A}$ and $\mathbf{B}$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

Multivariate normal (MVN) distribution

- Standard MVN
  - Random p-vector $\mathbf{Z} = (Z_1, \ldots, Z_p)^\top \sim \mathrm{MVN}(0, \mathbf{I}_p) \Leftrightarrow$ iid $Z_1 \ldots, Z_p \sim \mathcal{N}(0,1)$.
  - pdf of $\mathrm{MVN}(0, \mathbf{I}_p)$:

$$\begin{aligned} f_{\mathbf{Z}}(\boldsymbol{z}) &= \prod_{i=1}^{p} (2\pi)^{-1/2} \exp(-z_i^2/2) \\ &= (2\pi)^{-p/2} \exp(-\boldsymbol{z}^\top \boldsymbol{z}/2), \quad \boldsymbol{z} \in \mathbb{R}^p \end{aligned}$$

- In general
  - pdf of $\mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$:

$$f_{\mathbf{X}}(\boldsymbol{x}) = (2\pi)^{-p/2} (\det \boldsymbol{\Sigma})^{-1/2} \exp\{-(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})/2\}, \quad \boldsymbol{x} \in \mathbb{R}^p$$

  - $\mathbf{X} \sim \mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Rightarrow \mathbf{AX} + \boldsymbol{a} \sim \mathrm{MVN}(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{a}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$ for arbitrary $\mathbf{A} \in \mathbb{R}^{q \times p}$ and $\boldsymbol{a} \in \mathbb{R}^q$
  - $\mathbf{X} \sim \mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \mathbf{X} = \boldsymbol{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}$ with $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim \mathrm{MVN}(0, \mathbf{I}_p)$

Marginals of MVN

- Suppose p-vector $\mathbf{X} = (X_1, \ldots, X_p)^\top$ and q-vector $\mathbf{Y} = (Y_1, \ldots, Y_q)^\top$ are jointly normally distributed. Then, $\mathbf{X}$ and $\mathbf{Y}$ are independent $\Leftrightarrow \mathrm{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}_{p \times q}$.
- If $\mathbf{X}$ is of MVN, then its all margins are still of MVN. The inverse proposition does NOT hold.
  - Cautionary example: Let $Y = XZ$, where $X \sim \mathcal{N}(0,1)$; Z is independent of X with $\Pr(Z=1) = \Pr(Z=-1) = .5$. X and Y both turn out to be of standard normal, but they are not jointly normal.

Normal sampling theory (CB Sec. 5.3)

- Default identities for $X_1, \ldots, X_n$ iid as $\mathcal{N}(\mu, \sigma^2)$ (HMC Chp. 3)
  - $\sum_{i=1}^n X_i^2 \sim \chi^2(n)$ if iid $X_1, \ldots, X_n \sim \mathcal{N}(0,1)$.
  - $X/\sqrt{Y/n} \sim t(n)$ if $X \sim \mathcal{N}(0,1)$ and $Y \sim \chi^2(n)$ are independent.
  - $(X/m)/(Y/n) \sim F(m,n)$ if $X \sim \chi^2(m)$ and $Y \sim \chi^2(n)$ are independent.

- More identities for normal samples
  - $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ and $S^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$ are independent.
  - $n^{1/2}(\bar{X} - \mu)/\sigma \sim \mathcal{N}(0,1)$.
  - $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.
  - $n^{1/2}(\bar{X} - \mu)/S \sim t(n-1)$.