

Multivariate Distributions

Multinomial Distribution

- Suppose $0 \leq p_1, p_2, \dots, p_k$ are probabilities such that $\sum_{i=1}^k p_i = 1$.
- Consider a series of n identical trials where, on each trial, one can get exactly one of k possible outcomes o_1, \dots, o_k .
- Let $X_i = \#$ of trials resulting in outcome o_i .
- Then $X = (X_1, \dots, X_k)$ has a multinomial(n, p_1, \dots, p_k) distribution with joint pmf

$$\begin{aligned} & P(X_1 = x_1, \dots, X_\ell = x_\ell) \\ &= f(x_1, \dots, x_\ell) \\ &= \begin{cases} \frac{n!}{x_1!x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} & \text{integer } x_1, \dots, x_k \geq 0, \sum_{i=1}^k x_i = n \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- $\frac{n!}{x_1!x_2!\cdots x_k!}$ is the multinomial coefficient, counting number of ways/arrangements that n trials could result in x_1 outcomes o_1 , x_2 outcomes o_2, \dots, x_k outcomes o_k .
- Here individual marginal distributions are Binomial(n, p_i); conditionals of some given others are multinomial
- Example ($k = 4, n = 4$):

Multivariate Distributions

Dirichlet Distribution

- Suppose Y_1, \dots, Y_k are independent and $Y_i \sim \text{Gamma}(\alpha_i, 1)$, $i = 1, \dots, k$
- Let $X_i = Y_i / \sum_{j=1}^k Y_j$
- $0 \leq X_i \leq 1$
- $\sum_{i=1}^k X_i = 1$
- Then $X = (X_1, \dots, X_k)$ has a $\text{Dirichlet}(\alpha_1, \dots, \alpha_k)$ distribution
- Note that X_1, \dots, X_k are random variables which sum to 1
- Here individual marginal distributions are $\text{Beta}(\alpha_i, \sum_{j=1, j \neq i}^k \alpha_j)$; conditionals of some given others are scalar multiples of Dirichlets
- Example ($k = 4$): $\alpha_1 = 3, \alpha_2 = 2, \alpha_3 = 4, \alpha_4 = 1$

Note: $(X_1, X_2, \dots, X_k) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \dots, \alpha_k)$

$$\Rightarrow f_{X_1, \dots, X_k}(\alpha_1, \dots, \alpha_k) = \frac{\prod_{i=1}^k \alpha_i^{\alpha_i - 1}}{\prod_{i=1}^k \Gamma(\alpha_i)}$$

Multivariate Normal Distribution

Matrix-Vector Multivariate notation

- For a $p \times q$ matrix $\mathbf{A} = [A_{ij}]_{i=1,\dots,p; j=1,\dots,q}$ of random variables A_{ij} , $\mathbf{E}\mathbf{A}$ represents the matrix of expected values of components of \mathbf{A}
- For a $k \times 1$ random vector (r.v.), we can write

$$\mathbf{X} = \underbrace{(X_1, \dots, X_k)'}_{k \times 1} = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}_{k \times 1}$$

- The expected value of \mathbf{X} is

$$\mathbf{E}\mathbf{X} = \boldsymbol{\mu}_{\mathbf{X}} = \boldsymbol{\mu} = (\mathbf{E}X_1, \dots, \mathbf{E}X_k)'$$

- Definition:* The variance/covariance matrix of a $k \times 1$ r.v. $\mathbf{X} = (X_1, \dots, X_k)'$ is the $k \times k$ matrix

$$\text{Var}(\mathbf{X}) = \Sigma_{\mathbf{X}} = \Sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & \cdots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kk} \end{bmatrix}_{k \times k}$$

$\sigma_{ij} = \text{Cov}(X_i, X_j)$
 $\sigma_{ii} = \text{Var}(X_i)$

where $\sigma_{ij} = \text{Cov}(X_i, X_j)$

$$\mathbf{A}_{k \times k} \equiv \mathbf{B}_{k \times 1} \mathbf{B}'_{1 \times k}$$

- Result:

$$\text{Var}(\mathbf{X}) = \Sigma = \mathbf{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & \cdots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kk} \end{bmatrix}$$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{pmatrix} = (Y_1, \dots, Y_m)'$$

- If $\mathbf{Y} = (Y_1, \dots, Y_m)'$ is another r.v., then

$$X = (X_1, X_2, \dots, X_k)' = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}$$

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)'] = E[(X - \mu_X)(Y - \mu_Y)']_{k \times k \times m} = \begin{bmatrix} \text{Cov}(X_1, Y_1) & \dots & \text{Cov}(X_1, Y_m) \\ \vdots & \dots & \vdots \\ \text{Cov}(X_k, Y_1) & \dots & \text{Cov}(X_k, Y_m) \end{bmatrix}_{k \times m}$$

ij-Component is $E((X_i - EX_i)(Y_j - EY_j)) = E[X_i Y_j] - EX_i EY_j = \text{Cov}(X_i, Y_j)$

Transformation results: Let $\mathbf{X} = (X_1, \dots, X_k)'$ and $\mathbf{Y} = (Y_1, \dots, Y_m)'$ be r.v.s; let $A_{r \times k}$ and $B_{s \times m}$ be fixed $r \times k$ and $s \times m$ matrices; let $a_r = (a_1, \dots, a_r)'$ and $b_s = (a_1, \dots, a_s)'$ be fixed vectors; and let $d, c \in \mathbb{R}$ be constants.

$$1. E(a_r + A_{r \times k}X) = a_r + A_{r \times k}EX = a_r + A_{r \times k}\mu_X$$

$E(a + AX) = a + AEX$ (in one-dimension)

$$E(c + \sum_{i=1}^k a_i X_i) = E(c + a'_k X) = c + a'_k EX = c + a'_k \mu_X = c + \sum_{i=1}^k a_i \mu_i$$

$a'_k X = \sum_{i=1}^k a_i X_i$

$$2. \text{Var}(a_r + A_{r \times k}X) = A_{r \times k} \Sigma_X A'_{r \times k}$$

$\text{Var}(a + AX) = A^2 \text{Var}(X)$ in one-dimension.

$$\text{Var}\left(c + \sum_{i=1}^k a_i X_i\right) = \text{Var}(c + a'_k X) = a'_k \Sigma_X a_k = \sum_{i=1}^k \sum_{j=1}^k a_i a_j \text{Cov}(X_i, X_j)$$

$$3. \text{Cov}(a_r + A_{r \times k}X, b_s + B'_{s \times m}Y) = A_{r \times k} \text{Cov}(X, Y) B'_{s \times m}$$

$$\text{Var}\left(c + \sum_{i=1}^k a_i X_i, d + \sum_{j=1}^m b_j Y_j\right) = \text{Cov}(c + a'_k X, d + b'_m Y) = a'_k \text{Cov}(X, Y) b_m = \sum_{i=1}^k \sum_{j=1}^m a_i b_j \text{Cov}(X_i, Y_j)$$

Note: ② is coming from ③:

$$\text{Var}(a_r + A_{r \times k}X) = \text{Cov}(a_r + A_{r \times k}X, a_r + A_{r \times k}X) \stackrel{③}{=} A_{r \times k} \text{Cov}(X, X) A'_{r \times k} = A_{r \times k} \text{Var}(X) A'_{r \times k}$$

Definition: A (square) $k \times k$ matrix \mathbf{B} is

- ① non-singular if $\text{rank}(\mathbf{B}) = k$ or equivalently if $\det(\mathbf{B}) \neq 0$ or if \mathbf{B}^{-1} exists
- ② singular if $\text{rank}(\mathbf{B}) < k$ or equivalently if $\det(\mathbf{B}) = 0$ or if \mathbf{B}^{-1} fails to exist
- ③ **non-negative definite** if

$$\mathbf{a}'_k \mathbf{B} \mathbf{a}_k \geq 0 \quad \text{for any } \mathbf{a}_k = (a_1, \dots, a_k)' \in \mathbb{R}^k$$

- ④ **positive definite** if $\mathbf{a}'_k \mathbf{B} \mathbf{a}_k > 0$ for any non-zero $\mathbf{a}_k = (a_1, \dots, a_k)' \in \mathbb{R}^k$

(If \mathbf{B} is non-negative definite then $\det(\mathbf{B}) \geq 0$. A non-negative definite $k \times k$ matrix \mathbf{B} is positive definite iff \mathbf{B} has rank k iff \mathbf{B} is invertible, i.e., \mathbf{B}^{-1} exists iff $\det(\mathbf{B}) > 0$.)

Positive definite \Rightarrow non-negative definite

non-negative definite $\not\Rightarrow$ positive-definite

Non-negative definite + non-singular \Rightarrow Positive definite

Lemma: A $k \times k$ covariance matrix $\text{Var}(\mathbf{X}) = \Sigma$ is symmetric and non-negative definite. If Σ is not positive definite, then \mathbf{X} lies in a hyperplane $\{\mathbf{x}_k \in \mathbb{R}^k : \mathbf{a}'_k \mathbf{x}_k = b\} \subset \mathbb{R}^k$ for some non-zero $\mathbf{a}_k \in \mathbb{R}^k$ and some $b \in \mathbb{R}$ with probability 1.

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \quad t = \begin{pmatrix} t_1 \\ \vdots \\ t_k \end{pmatrix}$$

• Recall the mgf of $\mathbf{X} = (X_1, \dots, X_k)'$,

$$M_{\mathbf{X}}(\mathbf{t}) = M_{\mathbf{X}}(t_1, \dots, t_k) = Ee^{t_1 X_1 + \dots + t_k X_k} = Ee^{\sum_{i=1}^k t_i X_i} = Ee^{t' \mathbf{X}}$$

for $\mathbf{t} = (t_1, \dots, t_k)' \in \mathbb{R}^k$.

(The mgf of \mathbf{X} exists if the expected value exists for all \mathbf{t} in some open neighborhood of $\mathbf{0} = (0, \dots, 0)' \in \mathbb{R}^k$.)

- Recall also that if $M_{\mathbf{X}}(\mathbf{t}) = M_{\mathbf{Y}}(\mathbf{t})$ for all \mathbf{t} in some open neighborhood of $\mathbf{0}$ then \mathbf{X} and \mathbf{Y} have the same distribution.

Result: If $k \times 1$ random vector \mathbf{X} has mgf $M_{\mathbf{X}}(\mathbf{t})$, then for a given $\ell \times k$ matrix \mathbf{A} and given $\mathbf{b} \in \mathbb{R}^\ell$, the $\ell \times 1$ random vector $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ has mgf

$$M_{\mathbf{AX}+\mathbf{b}} = E(e^{t(\mathbf{AX}+\mathbf{b})}) = e^{t\mathbf{b}} M_{\mathbf{X}}(ta)$$

$$M_{\mathbf{Y}}(s) = e^{s'b} Ee^{As} \quad s = (s_1, \dots, s_\ell)' \in \mathbb{R}^\ell$$

$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_\ell \end{pmatrix}$ (provided $M_{\mathbf{Y}}(s)$ exists in an open neighborhood of $\mathbf{0} \in \mathbb{R}^\ell$)

$$M_{\mathbf{Y}}(s) = e^{s'b} M_{\mathbf{X}}(As)$$

$$M_{\mathbf{Y}}(s) = E e^{s'Y} = E e^{s'(\mathbf{AX}+\mathbf{b})}$$

$$= E e^{s'b} e^{s'AX}$$

$$= e^{s'b} E e^{s'AX}$$

$$= e^{s'b} E [e^{c'X}] = e^{s'b} M_{\mathbf{X}}(c)$$

$$= e^{s'b} M_{\mathbf{X}}(As)$$

Recall: If X_1, \dots, X_k are independent with mgfs $M_{X_i}(\cdot)$ then the $k \times 1$ random vector \mathbf{X} has mgf

* $M_{\mathbf{X}}(\mathbf{t}) = M_{\mathbf{X}}(t_1, \dots, t_k) = Ee^{\sum_{i=1}^k t_i X_i} = \prod_{i=1}^k M_{X_i}(t_i)$

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}$$