Stat 5101 Notes: Brand Name Distributions

Charles J. Geyer

September 18, 2006

1 Discrete Uniform Distribution

Symbol DiscreteUniform(n).

Type Discrete.

Rationale Equally likely outcomes.

Sample Space The interval 1, 2, . . ., n of the integers.

Probability Function

$$f(x) = \frac{1}{n}, \qquad x = 1, 2, \ldots, n$$

Moments

$$E(X) = \frac{n+1}{2}$$
$$\operatorname{var}(X) = \frac{n^2 - 1}{12}$$

2 Uniform Distribution

Symbol Uniform(a, b).

Type Continuous.

Rationale Continuous analog of the discrete uniform distribution.

Parameters Real numbers a and b with $a < b$.

Sample Space The interval (a, b) of the real numbers.

Probability Density Function

$$f(x) = \frac{1}{b-a}, \qquad a < x < b$$

Moments

$$\begin{aligned} E(X) &= \frac{a+b}{2} \\ \operatorname{var}(X) &= \frac{(b-a)^2}{12} \end{aligned}$$

Relation to Other Distributions Beta$(1,1) =$ Uniform$(0,1)$.

3 Bernoulli Distribution

Symbol Bernoulli(p).

Type Discrete.

Rationale Any zero-or-one-valued random variable.

Parameter Real number $0 \le p \le 1$.

Sample Space The two-element set $\{0,1\}$.

Probability Function

$$f(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$$

Moments

$$\begin{aligned} E(X) &= p \\ \operatorname{var}(X) &= p(1-p) \end{aligned}$$

Addition Rule If X_1, ..., X_k are i. i. d. Bernoulli(p) random variables, then $X_1 + \cdots + X_k$ is a Binomial(k,p) random variable.

Relation to Other Distributions Bernoulli$(p) =$ Binomial$(1,p)$.

4 Binomial Distribution

Symbol Binomial(n,p).

Type Discrete.

Rationale Sum of i. i. d. Bernoulli random variables.

Parameters Real number $0 \le p \le 1$. Integer $n \ge 1$.

Sample Space The interval 0, 1, ..., n of the integers.

Probability Function

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n$$

Moments

$$\begin{aligned} E(X) &= np \\ \operatorname{var}(X) &= np(1-p) \end{aligned}$$

Addition Rule If X_1, ..., X_k are independent random variables, X_i being Binomial(n_i, p) distributed, then $X_1 + \cdots + X_k$ is a Binomial$(n_1 + \cdots + n_k, p)$ random variable.

Normal Approximation If np and $n(1-p)$ are both large, then

$$\text{Binomial}(n, p) \approx \text{Normal}\big(np, np(1-p)\big)$$

Poisson Approximation If n is large but np is small, then

$$\text{Binomial}(n, p) \approx \text{Poisson}(np)$$

Theorem The fact that the probability function sums to one is equivalent to the **binomial theorem:** for any real numbers a and b

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a+b)^n.$$

Degeneracy If $p = 0$ the distribution is concentrated at 0. If $p = 1$ the distribution is concentrated at n.

Relation to Other Distributions Bernoulli(p) = Binomial$(1, p)$.

5 Hypergeometric Distribution

Symbol Hypergeometric(A, B, n).

Type Discrete.

Rationale Sample of size n without replacement from finite population of B zeros and A ones.

Sample Space The interval $\max(0, n - B)$, ..., $\min(n, A)$ of the integers.

Probability Function

$$f(x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}, \qquad x = \max(0, n - B), \ldots, \min(n, A)$$

Moments

$$\begin{aligned} E(X) &= np \\ \operatorname{var}(X) &= np(1-p) \cdot \frac{N-n}{N-1} \end{aligned}$$

where

$$\begin{aligned} p &= \frac{A}{A+B} \\ N &= A+B \end{aligned} \tag{5.1}$$

Binomial Approximation If n is small compared to either A or B, then

$$\text{Hypergeometric}(n, A, B) \approx \text{Binomial}(n, p)$$

where p is given by (5.1).

Normal Approximation If n is large, but small compared to either A or B, then

$$\text{Hypergeometric}(n, A, B) \approx \text{Normal}\big(np, np(1-p)\big)$$

where p is given by (5.1).

6 Poisson Distribution

Symbol Poisson(μ)

Type Discrete.

Rationale Counts in a Poisson process.

Parameter Real number $\mu > 0$.

Sample Space The non-negative integers 0, 1, ...

Probability Function

$$f(x) = \frac{\mu^x}{x!} e^{-\mu}, \qquad x = 0, 1, \ldots$$

Moments

$$E(X) = \mu$$
$$\mathrm{var}(X) = \mu$$

Addition Rule If X_1, ..., X_k are independent random variables, X_i being Poisson(μ_i) distributed, then $X_1 + \cdots + X_k$ is a Poisson($\mu_1 + \cdots + \mu_k$) random variable.

Normal Approximation If μ is large, then

$$\text{Poisson}(\mu) \approx \text{Normal}(\mu, \mu)$$

Theorem The fact that the probability function sums to one is equivalent to the Maclaurin series for the exponential function: for any real number x

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x.$$

7 Geometric Distribution

Symbol Geometric(p).

Type Discrete.

Rationales

- Discrete lifetime of object that does not age.
- Waiting time or interarrival time in sequence of i. i. d. Bernoulli trials.
- Inverse sampling.
- Discrete analog of the exponential distribution.

Parameter Real number $0 < p < 1$.

Sample Space The non-negative integers 0, 1, ...

Probability Function

$$f(x) = p(1-p)^x \qquad x = 0, 1, \ldots$$

Moments

$$E(X) = \frac{1-p}{p}$$
$$\operatorname{var}(X) = \frac{1-p}{p^2}$$

Addition Rule If X_1, ..., X_k are i. i. d. Geometric(p) random variables, then $X_1 + \cdots + X_k$ is a NegativeBinomial(k, p) random variable.

Theorem The fact that the probability function sums to one is equivalent to the geometric series: for any real number s such that $-1 < s < 1$

$$\sum_{k=0}^{\infty} s^k = \frac{1}{1-s}.$$

8 Negative Binomial Distribution

Symbol NegativeBinomial(r, p).

Type Discrete.

Rationale

- Sum of i. i. d. geometric random variables.
- Inverse sampling.

Parameters Real number $0 \le p \le 1$. Integer $r \ge 1$.

Sample Space The non-negative integers 0, 1, ...

Probability Function

$$f(x) = \binom{r+x-1}{x} p^r (1-p)^x, \qquad x = 0, 1, \ldots$$

Moments

$$E(X) = \frac{r(1-p)}{p}$$
$$\operatorname{var}(X) = \frac{r(1-p)}{p^2}$$

Addition Rule If X_1, ..., X_k are independent random variables, X_i being NegativeBinomial(r_i, p) distributed, then $X_1 + \cdots + X_k$ is a NegativeBinomial$(r_1 + \cdots + r_k, p)$ random variable.

Normal Approximation If $r(1-p)$ is large, then

$$\text{NegativeBinomial}(r, p) \approx \text{Normal}\left(\frac{r(1-p)}{p}, \frac{r(1-p)}{p^2}\right)$$

Extended Definition The definition makes sense for noninteger r if binomial coefficients are defined by

$$\binom{r}{k} = \frac{r \cdot (r-1) \cdots (r-k+1)}{k!}$$

which for integer r agrees with the standard definition.

Also

$$\binom{r+x-1}{x} = (-1)^x \binom{-r}{x} \tag{8.1}$$

which explains the name "negative binomial."

Theorem The fact that the probability function sums to one is equivalent to the **generalized binomial theorem:** for any real number s such that $-1 < s < 1$ and any real number m

$$\sum_{k=0}^{\infty} \binom{m}{k} s^k = (1+s)^m. \tag{8.2}$$

If m is a nonnegative integer, then $\binom{m}{k}$ is zero for $k > m$, and we get the ordinary binomial theorem.

Changing variables from m to $-m$ and from s to $-s$ and using (8.1) turns (8.2) into

$$\sum_{k=0}^{\infty} \binom{m+k-1}{k} s^k = \sum_{k=0}^{\infty} \binom{-m}{k} (-s)^k = (1-s)^{-m}$$

which has a more obvious relationship to the negative binomial density summing to one.

9 Normal Distribution

Symbol Normal(μ, σ^2).

Type Continuous.

Rationale Limiting distribution in the central limit theorem.

Parameters Real numbers μ and $\sigma^2 > 0$.

Sample Space The real numbers.

Probability Density Function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty$$

Moments

$$\begin{aligned} E(X) &= \mu \\ \operatorname{var}(X) &= \sigma^2 \\ E\{(X-\mu)^3\} &= 0 \\ E\{(X-\mu)^4\} &= 3\sigma^4 \end{aligned}$$

Linear Transformations If X is Normal(μ, σ^2) distributed, then $aX + b$ is Normal$(a\mu + b, a^2\sigma^2)$ distributed.

Addition Rule If $X_1, \ldots, X_k$ are independent random variables, X_i being Normal(μ_i, σ_i^2) distributed, then $X_1 + \cdots + X_k$ is a Normal$(\mu_1 + \cdots + \mu_k, \sigma_1^2 + \cdots + \sigma_k^2)$ random variable.

Theorem The fact that the probability density function integrates to one is equivalent to the integral

$$\int_{-\infty}^{\infty} e^{-z^2/2} \, dz = \sqrt{2\pi}$$

Relation to Other Distributions If Z is Normal$(0, 1)$ distributed, then Z^2 is Gamma$(\frac{1}{2}, \frac{1}{2})$ distributed.

10 Exponential Distribution

Symbol Exponential(λ).

Type Continuous.

Rationales

- Lifetime of object that does not age.
- Waiting time or interarrival time in Poisson process.
- Continuous analog of the geometric distribution.

Parameter Real number $\lambda > 0$.

Sample Space The interval $(0, \infty)$ of the real numbers.

Probability Density Function

$$f(x) = \lambda e^{-\lambda x}, \qquad 0 < x < \infty$$

Cumulative Distribution Function

$$F(x) = 1 - e^{-\lambda x}, \qquad 0 < x < \infty$$

Moments

$$\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \operatorname{var}(X) &= \frac{1}{\lambda^2} \end{aligned}$$

Addition Rule If X_1, ..., X_k are i. i. d. Exponential(λ) random variables, then $X_1 + \cdots + X_k$ is a Gamma(k, λ) random variable.

Relation to Other Distributions Exponential(λ) = Gamma($1, \lambda$).

11 Gamma Distribution

Symbol Gamma(α, λ).

Type Continuous.

Rationales

- Sum of i. i. d. exponential random variables.
- Conjugate prior for exponential, Poisson, or normal precision family.

Parameter Real numbers $\alpha > 0$ and $\lambda > 0$.

Sample Space The interval $(0, \infty)$ of the real numbers.

Probability Density Function

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad 0 < x < \infty$$

where $\Gamma(\alpha)$ is defined by (11.1) below.

Moments

$$E(X) = \frac{\alpha}{\lambda}$$
$$\operatorname{var}(X) = \frac{\alpha}{\lambda^2}$$

Addition Rule If $X_1, \ldots, X_k$ are independent random variables, X_i being Gamma(α_i, λ) distributed, then $X_1 + \cdots + X_k$ is a Gamma($\alpha_1 + \cdots + \alpha_k, \lambda$) random variable.

Normal Approximation If α is large, then

$$\text{Gamma}(\alpha, \lambda) \approx \text{Normal}\left(\frac{\alpha}{\lambda}, \frac{\alpha}{\lambda^2}\right)$$

Theorem The fact that the probability density function integrates to one is equivalent to the integral

$$\int_0^\infty x^{\alpha-1} e^{-\lambda x} \, dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}$$

the case $\lambda = 1$ is the definition of the *gamma function*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx \tag{11.1}$$

Relation to Other Distributions

- Exponential(λ) = Gamma($1, \lambda$).
- If X and Y are independent, X is $\Gamma(\alpha, \lambda)$ distributed and Y is $\Gamma(\beta, \lambda)$ distributed, then $X/(X + Y)$ is Beta(α, β) distributed.
- If Z is Normal($0, 1$) distributed, then Z^2 is Gamma($\frac{1}{2}, \frac{1}{2}$) distributed.

Facts About Gamma Functions Integration by parts in (11.1) establishes the **gamma function recursion formula**

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha), \qquad \alpha > 0 \tag{11.2}$$

The relationship between the Exponential(λ) and Gamma($1, \lambda$) distributions gives

$$\Gamma(1) = 1$$

and the relationship between the Normal($0, 1$) and Gamma($\frac{1}{2}, \frac{1}{2}$) distributions gives

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

Together with the recursion (11.2) these give for any positive integer n

$$\Gamma(n + 1) = n!$$

and

$$\Gamma(n + \tfrac{1}{2}) = \left(n - \tfrac{1}{2}\right)\left(n - \tfrac{3}{2}\right) \cdots \tfrac{3}{2} \cdot \tfrac{1}{2}\sqrt{\pi}$$

12 Beta Distribution

Symbol Beta(α, β).

Type Continuous.

Rationales

- Ratio of gamma random variables.
- Conjugate prior for binomial or negative binomial family.

Parameter Real numbers $\alpha > 0$ and $\beta > 0$.

Sample Space The interval $(0, 1)$ of the real numbers.

Probability Density Function

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha - 1}(1 - x)^{\beta - 1} \qquad 0 < x < 1$$

where $\Gamma(\alpha)$ is defined by (11.1) above.

Moments

$$E(X) = \frac{\alpha}{\alpha + \beta}$$

$$\operatorname{var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Theorem The fact that the probability density function integrates to one is equivalent to the integral

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Relation to Other Distributions

- If X and Y are independent, X is $\Gamma(\alpha, \lambda)$ distributed and Y is $\Gamma(\beta, \lambda)$ distributed, then $X/(X+Y)$ is $\text{Beta}(\alpha, \beta)$ distributed.
- $\text{Beta}(1, 1) = \text{Uniform}(0, 1)$.

13 Multinomial Distribution

Symbol $\text{Multinomial}(n, \mathbf{p})$

Type Discrete.

Rationale Multivariate analog of the binomial distribution.

Parameters Real vector $\mathbf{p}$ in the parameter space

$$\left\{ \mathbf{p} \in \mathbb{R}^k : 0 \le p_i,\ i = 1, \ldots, k,\ \text{and} \sum_{i=1}^k p_i = 1 \right\} \tag{13.1}$$

Sample Space The set of vectors with integer coordinates

$$S = \left\{ \mathbf{x} \in \mathbb{Z}^k : 0 \le x_i,\ i = 1, \ldots, k,\ \text{and} \sum_{i=1}^k x_i = n \right\} \tag{13.2}$$

Probability Function

$$f(\mathbf{x}) = \binom{n}{\mathbf{x}} \prod_{i=1}^k p_i^{x_i}, \qquad \mathbf{x} \in S$$

where

$$\binom{n}{\mathbf{x}} = \frac{n!}{\prod_{i=1}^k x_i!}$$

is called a *multinomial coefficient*.

Moments

$$\begin{aligned} E(X_i) &= np_i \\ \text{var}(X_i) &= np_i(1-p_i) \\ \text{cov}(X_i, X_j) &= -np_ip_j, \qquad i \ne j \end{aligned}$$

Moments (Vector Form)

$$E(\mathbf{X}) = n\mathbf{p}$$
$$\operatorname{var}(\mathbf{X}) = n\mathbf{M}$$

where

$$\mathbf{M} = \operatorname{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'$$

is the matrix with elements $m_{ij} = \operatorname{cov}(X_i, X_j)/n$.

Addition Rule If $\mathbf{X}_1$, ..., $\mathbf{X}_k$ are independent random vectors, $\mathbf{X}_i$ being Multinomial$(n_i, \mathbf{p})$ distributed, then $\mathbf{X}_1 + \cdots + \mathbf{X}_k$ is a Multinomial$(n_1 + \cdots + n_k, \mathbf{p})$ random variable.

Normal Approximation If n is large and $\mathbf{p}$ is not near the boundary of the parameter space (13.1), then

$$\text{Multinomial}(n, \mathbf{p}) \approx \text{Normal}(n\mathbf{p}, n\mathbf{M})$$

Theorem The fact that the probability function sums to one is equivalent to the **multinomial theorem:** for any vector $\mathbf{a}$ of real numbers

$$\sum_{\mathbf{x} \in S} \left[\binom{n}{\mathbf{x}} \prod_{i=1}^{k} a_i^{x_i} \right] = (a_1 + \cdots + a_k)^n$$

Degeneracy If there exists a vector $\mathbf{a}$ such that $\mathbf{M}\mathbf{a} = 0$, then $\operatorname{var}(\mathbf{a}'\mathbf{X}) = 0$.

In particular, the vector $\mathbf{u} = (1, 1, \ldots, 1)$ always satisfies $\mathbf{M}\mathbf{u} = 0$, so $\operatorname{var}(\mathbf{u}'\mathbf{X}) = 0$. This is obvious, since $\mathbf{u}'\mathbf{X} = \sum_{i=1}^{k} X_i = n$ by definition of the multinomial distribution, and the variance of a constant is zero. This means a multinomial random vector of dimension k is "really" of dimension no more than $k - 1$ because it is concentrated on a hyperplane containing the sample space (13.2).

Marginal Distributions Every univariate marginal is binomial

$$X_i \sim \text{Binomial}(n, p_i)$$

Not, strictly speaking marginals, but random vectors formed by collapsing categories are multinomial. If A_1, ..., A_m is a partition of the set $\{1, \ldots, k\}$ and

$$Y_j = \sum_{i \in A_j} X_i, \qquad j = 1, \ldots, m$$
$$q_j = \sum_{i \in A_j} p_i, \qquad j = 1, \ldots, m$$

then the random vector $\mathbf{Y}$ has a Multinomial$(n, \mathbf{q})$ distribution.

Conditional Distributions If $\{i_1, \ldots, i_m\}$ and $\{i_{m+1}, \ldots, i_k\}$ partition the set $\{1, \ldots, k\}$, then the conditional distribution of X_{i_1}, …, X_{i_m} given $X_{i_{m+1}}$, …, X_{i_k} is Multinomial$(n - X_{i_{m+1}} - \cdots - X_{i_k}, \mathbf{q})$, where the parameter vector $\mathbf{q}$ has components

$$q_j = \frac{p_{i_j}}{p_{i_1} + \cdots + p_{i_m}}, \qquad j = 1, \ldots, m$$

Relation to Other Distributions

- Each marginal of a multinomial is binomial.
- If X is Binomial(n, p), then the two-component vector $(X, n - X)$ is Multinomial$\bigl(n, (p, 1 - p)\bigr)$.

14 Bivariate Normal Distribution

Symbol See multivariate normal below.

Type Continuous.

Rationales See multivariate normal below.

Parameters Real vector $\boldsymbol{\mu}$ of dimension 2, real symmetric positive semi-definite matrix $\mathbf{M}$ of dimension 2×2 having the form

$$\mathbf{M} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

where $\sigma_1 > 0$, $\sigma_2 > 0$ and $-1 < \rho < +1$.

Sample Space The Euclidean space $\mathbb{R}^2$.

Probability Density Function

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{2\pi} \det(\mathbf{M})^{-1/2} \exp\left(-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'\mathbf{M}^{-1}(\mathbf{x} - \boldsymbol{\mu})'\right) \\ &= \frac{1}{2\pi\sqrt{1 - \rho^2}\sigma_1\sigma_2} \exp\Biggl(-\frac{1}{2(1 - \rho^2)}\Biggl[\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 \\ &\qquad -2\rho\left(\frac{x_1 - \mu_1}{\sigma_1}\right)\left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2\Biggr]\Biggr), \qquad \mathbf{x} \in \mathbb{R}^2 \end{aligned}$$

Moments

$$\begin{aligned} E(X_i) &= \mu_i, \qquad i = 1, 2 \\ \operatorname{var}(X_i) &= \sigma_i^2, \qquad i = 1, 2 \\ \operatorname{cov}(X_1, X_2) &= \rho\sigma_1\sigma_2 \\ \operatorname{cor}(X_1, X_2) &= \rho \end{aligned}$$

Moments (Vector Form)

$$\begin{aligned} E(\mathbf{X}) &= \boldsymbol{\mu} \\ \operatorname{var}(\mathbf{X}) &= \mathbf{M} \end{aligned}$$

Linear Transformations See multivariate normal below.

Addition Rule See multivariate normal below.

Marginal Distributions X_i is $\text{Normal}(\mu_i, \sigma_i^2)$ distributed, $i = 1, 2$.

Conditional Distributions The conditional distribution of X_2 given X_1 is

$$\text{Normal}\Big(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), (1 - \rho^2)\sigma_2^2\Big)$$

15 Multivariate Normal Distribution

Symbol $\text{Normal}(\boldsymbol{\mu}, \mathbf{M})$

Type Continuous.

Rationales

- Multivariate analog of the univariate normal distribution.
- Limiting distribution in the multivariate central limit theorem.

Parameters Real vector $\boldsymbol{\mu}$ of dimension k, real symmetric positive semi-definite matrix $\mathbf{M}$ of dimension $k \times k$.

Sample Space The Euclidean space $\mathbb{R}^k$.

Probability Density Function If $\mathbf{M}$ is (strictly) positive definite,

$$f(\mathbf{x}) = (2\pi)^{-k/2} \det(\mathbf{M})^{-1/2} \exp\left(-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'\mathbf{M}^{-1}(\mathbf{x} - \boldsymbol{\mu})'\right), \qquad \mathbf{x} \in \mathbb{R}^k$$

Otherwise there is no density ($\mathbf{X}$ is concentrated on a hyperplane).

Moments (Vector Form)

$$E(\mathbf{X}) = \boldsymbol{\mu}$$
$$\operatorname{var}(\mathbf{X}) = \mathbf{M}$$

Linear Transformations If $\mathbf{X}$ is Normal$(\boldsymbol{\mu}, \mathbf{M})$ distributed, then $\mathbf{AX} + \mathbf{b}$, where $\mathbf{A}$ is a constant matrix and $\mathbf{b}$ is a constant vector of dimensions such that the matrix multiplication and vector addition make sense, is Normal$(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{AMA}')$ distributed.

Addition Rule If $\mathbf{X}_1$, …, $\mathbf{X}_k$ are independent random vectors, $\mathbf{X}_i$ being Normal$(\boldsymbol{\mu}_i, \mathbf{M}_i)$ distributed, then $\mathbf{X}_1 + \cdots + \mathbf{X}_k$ is a Normal$(\boldsymbol{\mu}_1 + \cdots + \boldsymbol{\mu}_k, \mathbf{M}_1 + \cdots + \mathbf{M}_k)$ random variable.

Degeneracy If there exists a vector $\mathbf{a}$ such that $\mathbf{Ma} = 0$, then $\operatorname{var}(\mathbf{a}'\mathbf{X}) = 0$.

Partitioned Vectors and Matrices The random vector and parameters are written in *partitioned form*

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \tag{15.1a}$$

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \tag{15.1b}$$

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_2 \end{pmatrix} \tag{15.1c}$$

when $\mathbf{X}_1$ consists of the first r elements of $\mathbf{X}$ and $\mathbf{X}_2$ of the other $k - r$ elements and similarly for $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$.

Marginal Distributions Every marginal of a multivariate normal is normal (univariate or multivariate as the case may be). In partitioned form, the (marginal) distribution of $\mathbf{X}_1$ is Normal$(\boldsymbol{\mu}_1, \mathbf{M}_{11})$.

Conditional Distributions Every conditional of a multivariate normal is normal (univariate or multivariate as the case may be). In partitioned form, the conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2$ is

$$\text{Normal}(\boldsymbol{\mu}_1 + \mathbf{M}_{12}\mathbf{M}_{22}^{-}[\mathbf{X}_2 - \boldsymbol{\mu}_2], \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-}\mathbf{M}_{21})$$

where the notation $\mathbf{M}_{22}^{-}$ denotes the inverse of the matrix $\mathbf{M}_{22}^{-}$ if the matrix is invertable and otherwise any generalized inverse.