Chapter 2. Multivariate Distributions
2.6. Extension to Several Random Variables—Proofs of Theorems
Table of contents

1. Theorem 2.6.1
2. Theorem 2.6.2
3. Theorem 2.6.3
4. Corollary 2.6.A
Theorem 2.6.1. Suppose $X_1, X_2, \ldots, X_n$ are $n$ mutually independent random variables. Suppose the Moment generating function for $X_i$ is $M_i(t)$ for $-j_1 < t < h_i$ where $h_i > 0$, for $i = 1, 2, \ldots, n$. Let $T = \sum_{i=1}^{n} k_i X_i$ where $k_1, k_2, \ldots, k_n$ are constants. Then $T$ has the moment generating function given by 

$$M_T(t) = \prod_{i=1}^{n} M_i(k_i t) \quad \text{for} \quad -\min_{1 \leq i \leq n}\{h_i\} \leq t \leq \min_{1 \leq i \leq n}\{h_i\}.$$ 

Proof. Assume $t$ is in the interval $(-\min_{1 \leq i \leq n}\{h_i\}, \min_{1 \leq i \leq n}\{h_i\})$. Then

$$M_T(t) = E \left[ \exp \left( \sum_{i=1}^{n} tk_i X_i \right) \right] = E \left[ \prod_{i=1}^{n} e^{ik_i X_i} \right]$$

$$= \prod_{i=1}^{n} E \left[ e^{tk_i X_i} \right] \text{ by the mutual independence}$$

$$= \prod_{i=1}^{n} M_i(k_i t).$$
Theorem 2.6.1. Suppose $X_1, X_2, \ldots, X_n$ are $n$ mutually independent random variables. Suppose the Moment generating function for $X_i$ is $M_i(t)$ for $-j_1 < t < h_i$ where $h_i > 0$, for $i = 1, 2, \ldots, n$. Let $T = \sum_{i=1}^n k_i X_i$ where $k_1, k_2, \ldots, k_n$ are constants. Then $T$ has the moment generating function given by

$$M_T(i) = \prod_{i=1}^n M_i(k_i t) \quad \text{for} \quad - \min_{1 \leq i \leq n} \{h_i\} \leq t \leq \min_{1 \leq i \leq n} \{h_i\}.$$ 

**Proof.** Assume $t$ is in the interval $(- \min_{1 \leq i \leq n} \{h_i\}, \min_{1 \leq i \leq n} \{h_i\})$. Then

$$M_T(t) = E \left[ \exp \left( \sum_{i=1}^n t k_i X_i \right) \right] = E \left[ \prod_{i=1}^n e^{t k_i X_i} \right]$$

$$= \prod_{i=1}^n E \left[ e^{t k_i X_i} \right] \quad \text{by the mutual independence}$$

$$= \prod_{i=1}^n M_i(k_i t).$$
Theorem 2.6.2

**Theorem 2.6.2.** Let $V$ and $W$ be $m \times n$ matrices of random variables, let $A$ and $C$ be $k \times m$ matrices of constants, and let $B$ be an $n \times \ell$ matrix of constants. Then $E[AV + CW] = AE[V] + CE[W]$ and $E[AWB] = AE[E]B$; that is, $E$ is a linear operator on matrices of random variables.

**Proof.** Since $E$ is a linear operator on random variable by Theorem 2.1.1, then the $(i, j)$ component of $E[AV + CW]$ is

$$E \left[ \sum_{s=1}^{m} a_{is} V_{sj} + \sum_{s=1}^{m} c_{is} W_{sj} \right] = \sum_{s=1}^{m} a_{is} E[V_{sj}] + \sum_{s=1}^{m} c_{is} E[W_{sj}]$$

and the first claim holds.
Theorem 2.6.2

Theorem 2.6.2. Let $V$ and $W$ be $m \times n$ matrices of random variables, let $A$ and $C$ be $k \times m$ matrices of constants, and let $B$ be an $n \times \ell$ matrix of constants. Then $E[AV + CW] = AE[V] + CE[W]$ and $E[AWB] = AE[E]B$; that is, $E$ is a linear operator on matrices of random variables.

Proof. Since $E$ is a linear operator on random variable by Theorem 2.1.1, then the $(i, j)$ component of $E[AV + CW]$ is

$$E \left[ \sum_{s=1}^{m} a_{is} V_{sj} + \sum_{s=1}^{m} c_{is} W_{sj} \right] = \sum_{s=1}^{m} a_{is} E[V_{sj}] + \sum_{s=1}^{m} c_{is} E[W_{sj}]$$

and the first claim holds.

Next, the $(i, p)$ entry of $AW$ (an $k \times m$ matrix) is $\sum_{s=1}^{m} a_{is} W_{sp}$ and the $(i, j)$ entry of $AWB$ (an $k \times \ell$ matrix) is $\sum_{p=1}^{n} (\sum_{s=1}^{m} a_{is} W_{sp}) b_{pj}$. 
**Theorem 2.6.2.** Let $\mathbf{V}$ and $\mathbf{W}$ be $m \times n$ matrices of random variables, let $\mathbf{A}$ and $\mathbf{C}$ be $k \times m$ matrices of constants, and let $\mathbf{B}$ be an $n \times \ell$ matrix of constants. Then $E[\mathbf{AV} + \mathbf{CW}] = \mathbf{A} E[\mathbf{V}] + \mathbf{C} E[\mathbf{W}]$ and $E[\mathbf{AWB}] = \mathbf{A} E[\mathbf{E}] \mathbf{B}$; that is, $E$ is a linear operator on matrices of random variables.

**Proof.** Since $E$ is a linear operator on random variable by Theorem 2.1.1, then the $(i, j)$ component of $E[\mathbf{AV} + \mathbf{CW}]$ is

$$E \left[ \sum_{s=1}^{m} a_{is} V_{sj} + \sum_{s=1}^{m} c_{is} W_{sj} \right] = \sum_{s=1}^{m} a_{is} E[V_{sj}] + \sum_{s=1}^{m} c_{is} E[W_{sj}]$$

and the first claim holds.

Next, the $(i, p)$ entry of $\mathbf{AW}$ (an $k \times m$ matrix) is $\sum_{s=1}^{m} a_{is} W_{sp}$ and the $(i, j)$ entry of $\mathbf{AWB}$ (an $k \times \ell$ matrix) is $\sum_{p=1}^{n} (\sum_{s=1}^{m} a_{is} W_{sp}) b_{pj}$. 
Theorem 2.6.2 (continued)

**Proof (continued).** Since $E$ is a linear operator on random variables by Theorem 2.1.1, then the $(i, j)$ entry of $E[\mathbf{AWB}]$ is

$$E \left[ \sum_{p=1}^{n} \left( \sum_{s=1}^{m} a_{is} W_{sp} \right) b_{pj} \right] = \sum_{p=1}^{m} E \left[ \sum_{s=1}^{m} a_{is} W_{sp} b_{pj} \right]$$

$$= \sum_{p=1}^{m} E \left[ \sum_{s=1}^{m} a_{is} W_{sp} \right] b_{pj} = \sum_{p=1}^{m} \left( \sum_{s=1}^{m} a_{is} E[W_{sp}] \right) b_{pj},$$

and this is the $(i, j)$ entry of $\mathbf{A}E[\mathbf{W}]\mathbf{B}$, so the second claim holds.
Theorem 2.6.3

Theorem 2.6.3. Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)' = (X_1, X_2, \ldots, X_n)^T$ be an $n$-dimensional random vector, such that $\sigma_i^2 = \sigma_{ii} = \text{Var}(X_i) < \infty$. Let $A$ be an $m \times n$ matrix of constants. Then $\text{Cov}(\mathbf{X}) = E[\mathbf{XX}'] = \mu\mu'$ and $\text{Cov}(A\mathbf{X}) = A\text{Cov}(\mathbf{X})A'$.

Proof. First,

$$\text{Cov}(\mathbf{X}) = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)']$$ by definition
$$= E[\mathbf{XX}' - \mu \mathbf{X}' - \mathbf{X} \mu' + \mu \mu']$$
$$= E[\mathbf{XX}'] - \mu E[\mathbf{X}'] = E[\mathbf{X}]\mu' + E[\mu\mu']$$ by Theorem 2.6.2
$$= E[\mathbf{XX}'] - \mu \mu' - \mu \mu' + \mu \mu'$$
$$= E[\mathbf{XX}'] - \mu \mu'$$,

as claimed.
**Theorem 2.6.3**

**Theorem 2.6.3.** Let \( X = (X_1, X_2, \ldots, X_n)' = (X_1, X_2, \ldots, X_n)^T \) be an \( n \)-dimensional random vector, such that \( \sigma_i^2 = \sigma_{ii} = \text{Var}(X_i) < \infty \). Let \( A \) be an \( m \times n \) matrix of constants. Then \( \text{Cov}(X) = E[XX'] = \mu \mu' \) and \( \text{Cov}(AX) = A \text{Cov}(X) A' \).

**Proof.** First,

\[
\text{Cov}(X) = E[(X - \mu)(X - \mu)'] \text{ by definition}
\]
\[
= E[XX' - \mu X' - X \mu' + \mu \mu']
\]
\[
= E[XX'] - \mu E[X'] = E[X] \mu' + E[\mu \mu'] \text{ by Theorem 2.6.2}
\]
\[
= E[XX'] - \mu \mu' - \mu \mu' + \mu \mu'
\]
\[
= E[XX'] - \mu \mu',
\]

as claimed.
**Theorem 2.6.3** (continued)

**Theorem 2.6.3.** Let $X = (X_1, X_2, \ldots, X_n)' = (X_1, X_2, \ldots, X_n)^T$ be an $n$-dimensional random vector, such that $\sigma^2_i = \sigma_{ii} = \text{Var}(X_i) < \infty$. Let $A$ be an $m \times n$ matrix of constants. Then $\text{Cov}(X) = E[XX'] = \mu\mu'$ and $\text{Cov}(AX) = A\text{Cov}(X)A'$.

**Proof (continued).** Next, by Theorem 2.6.2, $E[AX] = AE[X] = A\mu$ and

$$\text{Cov}(AX) = E[(AA - A\mu)(AX - A\mu)'] \text{ by definition}$$
$$= E[(AA - A\mu)(X'A' - \mu'A')]$$
$$\text{since } (AB)' = (AB)^T = B^T A^T = B'A'$$
$$= E[AXX'A' - A\mu X'A' - AX\mu'A' + A\mu\mu'A']$$
$$= \text{Cov}(X)A' \text{ by the first result.}$$
Corollary 2.6.A

**Corollary 2.6.A.** All variance-covariance matrices are positive semi-definite.

**Proof.** Let $\mathbf{X}$ be a random (column) vector of $n$ random variables and let $\mathbf{a}$ be a constant $n \times 1$ vector. Then $\mathbf{Y} = \mathbf{a}'\mathbf{X}$ is a random variable (a linear combination of the components of $\mathbf{X}$) and so has a nonnegative variance. That is,

$$0 \leq \text{Var}(\mathbf{Y}) = \text{Var}(\mathbf{a}'\mathbf{X})$$

$$= E[(\mathbf{a}'\mathbf{X} - E[\mathbf{a}'\mathbf{X}])^2] \text{ by Definition 1.9.2}$$

$$= \text{Cov}(\mathbf{a}'\mathbf{X}) \text{ since Cov}(\mathbf{Y}) = \text{Var}(\mathbf{Y}) \text{ for a single random variable}$$

$$= \mathbf{a}'\text{Cov}(\mathbf{X})\mathbf{a} \text{ by Theorem 2.6.3.}$$

So Cov($\mathbf{X}$) is a positive semi-definite matrix, as claimed.
Corollary 2.6.A

Corollary 2.6.A. All variance-covariance matrices are positive semi-definite.

Proof. Let $X$ be a random (column) vector of $n$ random variables and let $a$ be a constant $n \times 1$ vector. Then $Y = a'X$ is a random variable (a linear combination of the components of $X$) and so has a nonnegative variance. That is,

\[
0 \leq \text{Var}(Y) = \text{Var}(a'X) \\
= E[(a'X - E[a'X])^2] \text{ by Definition 1.9.2} \\
= \text{Cov}(a'X) \text{ since Cov}(Y) = \text{Var}(Y) \text{ for a single random variable} \\
= a'\text{Cov}(X)a \text{ by Theorem 2.6.3}.
\]

So Cov($X$) is a positive semi-definite matrix, as claimed. \qed