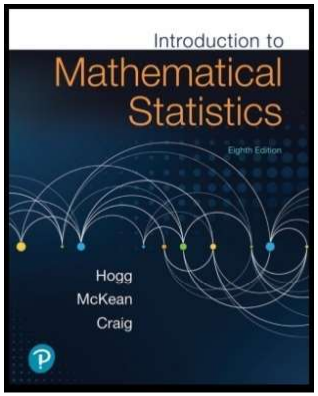
Mathematical Statistics 1

Chapter 2. Multivariate Distributions

2.8. Linear Combinations of Random Variables—Proofs of Theorems



Theorem 2.8.1

Theorem 2.8.1. Let $X_1, X_2, \ldots, X_n$ be random variables and define $T = \sum_{i=1}^n a_i X_i$. Suppose $E[X_i] = \mu_i$ for $i = 1, 2, \ldots, n$. Then $E[T] = \sum_{i=1}^n a_i \mu_i$.

Proof. By Theorem 2.1.1 (and induction) E is linear so

$$E[T] = E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n E[a_i X_i] = \sum_{i=1}^n a_i E[X_i] = \sum_{i=1}^n a_i \mu_i,$$

as claimed. □

Theorem 2.8.2

Theorem 2.8.2. Let $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m$ be random variables and define $T = \sum_{i=1}^n a_i X_i$ and $W = \sum_{j=1}^m b_j Y_j$. If $E[X_i^2] < \infty$ and $E[Y_j^2] < \infty$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$, then

$$\mathrm{Cov}(T, W) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \,\mathrm{Cov}(X_i, Y_j).$$

Proof. We have by the definition of covariance (Definition 2.5.1) that

$$\begin{aligned}
\mathrm{Cov}(T, W) &= E[(T - \mu_T)(W - \mu_W)] \text{ where } E[T] = \mu_T \text{ is the} \\
&\quad \text{mean of } T \text{ and } E[W] = \mu_W \text{ is the mean of } W \\
&= E\left[\left(\sum_{i=1}^n a_i X_i = \sum_{i=1}^n a_i E[X_i]\right)\left(\sum_{j=1}^m b_j Y_j - \sum_{j=1}^m b_j E[Y_j]\right)\right] \\
&= E\left[\left(\sum_{i=1}^n (a_i X_i - a_i E[X_i])\right)\left(\sum_{j=1}^m (b_j Y_j - b_j E[Y_j])\right)\right]
\end{aligned}$$

Theorem 2.8.2 (continued 1)

Proof (continued).

$$\begin{aligned}
\mathrm{Cov}(T, W) &= E\left[\sum_{i=1}^n \sum_{j=1}^m (a_i X_i - a_i E[X_i])(b_j Y_j - b_j E[Y_j])\right] \\
&= \sum_{i=1}^n \sum_{j=1}^m E[(a_i X_i - a_i E[X_i])(b_j Y_j - b_j E[Y_j])] \\
&\quad \text{since } E \text{ is linear by Theorem 2.1.1} \\
&= \sum_{i=1}^n \sum_{j=1}^m E[a_i b_j X_i Y_j - a_i b_j E[X_i] Y_j \\
&\quad - a_i b_j E[X_i] Y_j + a_i b_j E[X_i] E[Y_j]) \\
&= \sum_{i=1}^n \sum_{j=1}^m (a_i b_j E[X_i Y_j] - a_i b_j E[X_i] E[Y_j] \\
&\quad - a_i b_j E[X_i] E[Y_j] + a_i b_j E[X_i] E[Y_j]) \text{ by Theorem 2.1.1}
\end{aligned}$$

Theorem 2.8.2 (continued 2)

Proof (continued). ...

$$\begin{aligned}\operatorname{Cov}(T,W) &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j (E[X_iY_j] - 2E[X_i]E[Y_j] + E[X_i]E[Y_j]) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j E[(X_i - E[X_i])(Y_j - E[Y_j])] \text{ by Theorem 2.1.1} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \operatorname{Cov}(X_i, Y_j) \text{ by Definition 2.5.1,}\end{aligned}$$

as claimed. □

Corollary 2.8.1

Corollary 2.8.1. Let $X_1, X_2, \ldots, X_n$ be random variables and define $T = \sum_{i=1}^n a_i X_i$. Provided $E[X_i^2] < \infty$ for $i = 1, 2, \ldots, n$, then

$$\operatorname{Var}(T) = \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + s \sum_{i<j} a_i a_j \operatorname{Cov}(X_i, X_j).$$

Proof. By the definition of variance (Definition 1.9.2) we have

$$\begin{aligned}\operatorname{Var}(T) &= E[(T - E[T])^2) = E[(T - E[T])(T - E[T])] \\ &= \operatorname{Cov}(T, T) \text{ by Definition 2.5.1} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i a_j \operatorname{Cov}(X_i, X_i) + \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{m} a_i a_j \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{m} a_i a_j \operatorname{Cov}(X_i, X_j)\end{aligned}$$

since $\operatorname{Var}(X_i) = \operatorname{Cov}(X_i, X_i)$ and $\operatorname{Cov}(X_i, X_j) > \operatorname{Cov}(X_j, X_i)$

Corollary 2.8.1 (continued)

Corollary 2.8.1. Let $X_1, X_2, \ldots, X_n$ be random variables and define $T = \sum_{i=1}^n a_i X_i$. Provided $E[X_i^2] < \infty$ for $i = 1, 2, \ldots, n$, then

$$\operatorname{Var}(T) = \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + s \sum_{i<j} a_i a_j \operatorname{Cov}(X_i, X_j).$$

Proof (continued). ...

$$\begin{aligned}\operatorname{Var}(T) &= \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{m} a_i a_j \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + 2\sum_{i<j} a_i a_j \operatorname{Cov}(X_i, X_j),\end{aligned}$$

as claimed. □

Corollary 2.8.2

Corollary 2.8.2. Let $X_1, X_2, \ldots, X_n$ be independent random variables and define $T = \sum_{i=1}^n a_i X_i$. With $\operatorname{Var}(X_i) = \sigma_i^2$ for $i = 1, 2, \ldots, n$ we have $\operatorname{Var}(T) = \sum_{i=1}^n a_i^2 \sigma_i^2$.

Proof. Since X_i and X_j are independent for $i \neq j$ then by Theorem 2.5.2 $\operatorname{Cov}(X_i, X_j) = 0$ for $i \neq j$. So by Corollary 2.8.1,

$$\operatorname{Var}(T) = \operatorname{Cor}(T, T) = \sum_{i=1}^{n} a_i^2 \operatorname{Var}(X_i) + 2\sum_{i<j} a_i a_j \operatorname{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} a_i \sigma_i^2 + 0 = \sum_{i=1}^{n} a_i \sigma_i^2,$$

as claimed. □

Theorem 2.8.A

Theorem 2.8.A. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with common mean μ and variance σ^2. We have

$$E[\overline{X}] = \mu, \ \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}, \ S^2 = \frac{\sum_{i=1}^n X_i^2 - n\overline{X}^2}{n-1}, \text{ and } E[S^2] = \sigma^2.$$

Proof. We have $\overline{X} = \frac{\sum_{i=1}^n X_i}{n} = \sum_{i=1}^n \frac{X_i}{n}$ by definition, so by Theorem 2.8.1

$$E[\overline{X}] = \sum_{i=1}^n \frac{1}{n} E[X_i] = \sum_{i=1}^n \frac{\mu}{n} = \mu.$$

By Corollary 2.8.2, $\mathrm{Var}(\overline{X}) = \sum_{i=1}^n \frac{1}{n^2}\sigma^2 = \frac{\sigma^2}{n}$. The proof that $S^2 = \frac{\sum_{i=1}^n X_i^2 - n\overline{X}^2}{n-1}$ is to be given in Exercise 2.8.1.

Theorem 2.8.A (continued)

Proof (continued). Finally,

$$\begin{aligned} E[S^2] &= E\left[\frac{\sum_{i=1}^n X_i^2 - n\overline{X}^2}{n-1}\right] \\ &= \frac{\sum_{i=1}^n E[X_i^2] - nE[\overline{X}^2]}{n-1} \text{ by Theorem 2.1.1} \\ &= \frac{n(\sigma^2 - \mu^2) - nE[\overline{X}^2]}{n-1} \text{ since } E[X^2] = \sigma^2 + \mu^2 \text{ by Note 1.9.A} \\ &= \frac{n(\sigma^2 + \mu^2) - n(\sigma^2/n + \mu^2)}{n-1} \text{ since } E(\overline{X}^2) = \mathrm{Var}(\overline{X})^2 + E(\overline{X})^2 \\ &\qquad\qquad = \frac{\sigma^2}{n} + \mu^2 \\ &= \frac{(n-1)\sigma^2}{n-1} = \sigma^2, \end{aligned}$$

as claimed. □