General Random Variables

1 Expectation

The expectation EX of a general random variable X is defined in section 1.2 of the text. The properties of expectation for a general random variable are not proved. We will assume the usual facts:

(i)

$$E(aX + bY) = aEX + bEY.$$

(ii)

$$EX \leq EY \text{ if } X \leq Y \text{ a.s.}$$

(iii) If $X \geq 0$ and $EX = 0$ then $X = 0$ a.s.

2 Independence

Let X be a random variable taking values in a set S. Let Y be a random variable taking values in a set T. We say that X and Y are *independent* if

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$$

for every $A \subset S$ and every $B \subset T$.

Technical point For a general random variable, we actually cannot consider *all* subsets A of S and *all* subsets B of T. We are only allowed to consider *measurable* subsets. Fortunately, we will never encounter any sets which are not measurable, so this technicality will be *ignored* from now on. In particular, we will not state the definition of a measurable set or a measurable function. We just assume that any set or function we talk about is measurable.

More generally, if we are given a random variable X_j taking values in a set S_j, for $j = 1, \ldots, k$, we say that $X_1, \ldots, X_k$ is an independent sequence if

$$P(X_1 \in A_1, \ldots, X_k \in A_k) = P(X_1 \in A_1) \ldots P(X_k \in A_k)$$

for all sets $A_j \subset S_j$, $j = 1, \ldots, k$.

Lemma 1 (Functions of independent are independent)
Suppose that $X_1, \ldots, X_{m+n}$ is an independent sequence of random variables taking values in sets $S_1, \ldots, S_{m+n}$, respectively. Let f be a function from $S_1 \times \ldots \times S_n$ to some set S and let g be a function from $S_{n+1} \times \ldots \times S_{n+m}$ to some set T. Then the random variables $f(X_1, \ldots, X_n)$ and $g(X_{n+1}, \ldots, X_{n+m})$ are independent.

(A corresponding statement holds for more than two functions.)

We won't prove Lemma 1, but we use it sometimes.

Most of the random variables we deal with now take values in $\mathbb{R}^d$ for various values of d. Our text expects that we know the facts which follow next.

Lemma 2 (Expectations and independence) *Let $X_1, \ldots, X_n$ be an independent sequence of real-valued random variables having finite expectation. Then $X_1 \ldots X_n$ has finite expectation, and*

$$EX_1 \ldots X_n = (EX_1) \ldots (EX_n).$$

If the random variables $X1, \ldots, X_n$ have finite second moments then

$$Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n).$$

Lemma 3 (Characteristic functions and independence) *Let X be an $\mathbb{R}^d$-valued random variable and let Y be an $\mathbb{R}^{d_1}$-valued random variable. The following statements are equivalent:*

(i) X and Y are independent.

(ii)

$$\phi_{(X,Y)}(t_1, \ldots, t_{d+d_1}) = \phi_X(t_1, \ldots, t_d)\, \phi_Y(t_{d+1}, \ldots, t_{d+d_1}).$$

3 Covariance facts

Discussed in class.

Lemma 4 (nonsingular Σ test) *Let X be a mean zero $\mathbb{R}^d$-valued random variable with second moments finite. The following statements are equivalent.*

(i) $X_1, \ldots, X_d$ are independent (where "0" means "0 a.s.").

(ii) $Var(c_1X_1 + \ldots + c_dX_d) = 0$ only if $c_1 = \ldots = c_d$.

(iii) $\Sigma_{X,X}$ is nonsingular.

The next lemma contains the key argument for Problem 4.3.8.

Lemma 5 (Covariance projection formula) *Let Y be a $\mathbb{R}^{d_1}$-valued random variable with second moments finite. Let X be a $\mathbb{R}^d$-valued random variable with second moments finite. Assume that $\Sigma_{Y,Y}$ is nonsingular.*

Let

$$Z = \Sigma_{X,Y}\Sigma_{Y,Y}^{-1}Y.$$

Then $X - Z$ and Y are uncorrelated, that is,

$$\Sigma_{X-Z,Y} = 0.$$

Let $\mathcal{L}$ be the span of $Y_1, \ldots, Y_{d_1}$. If X and Y are mean zero then

$$Z_j = \Pi_{\mathcal{L}}\, X_j$$

for each j.

Proof By covariance properties from page 115 we have

$$\begin{aligned} \Sigma_{X-Z,Y} &= \Sigma_{X,Y} - \Sigma_{Z,Y} \\ &= \Sigma_{X,Y} - \Sigma_{X,Y}\Sigma_{Y,Y}^{-1}\Sigma_{Y,Y} = 0. \end{aligned}$$

Since $\Sigma_{X-Z,Y} = 0$ we have $\mathrm{Cov}\,(X_j - Z_j, Y_k) = 0$ for every j, k and hence $\mathrm{Cov}\,(X_j - Z_j, W) = 0$ for every $W \in \mathcal{L}$.

Now assume that X and Y are mean zero. Then for each j we have $X_j - Z_j \perp \mathcal{L}$, and hence by definition $Z_j = \Pi_{\mathcal{L}}\, X_j$.

This proves the lemma.

4 1-dimensional Gaussians

A real-valued random variable is Gaussian if it is a constant, or if it has a normal distribution.

This definition agrees with Definition 4.2.18, although it doesn't talk about characteristic functions.

Note equation (4) tells us what the characteristic function of a real-valued Gaussian looks like. This is easy to check but we will just accept it and use it.

5 n-dimensional Gaussians

According to Definition 4.2.18, an $\mathbb{R}^n$-valued random variable $X = (X_1, \ldots, X_n)$ is Gaussian if every real-valued random variable $Z = c_1X_1 + \ldots + c_nX_n$ is Gaussian.

Problem 4.2.20 tells us what the characteristic function of an $\mathbb{R}^n$-valued Gaussian looks like.

6 Generalizing the Gaussian definition

Any set G of random variables is called a Gaussian set if for any random variables $Y_1, \ldots, Y_k \in G$, and any constants $c_1, \ldots, c_k$, the random variable $Z = c_1 Y_1 + \ldots + c_k Y_k$ is Gaussian.

It turns out that this general definition makes some statements clearer.

Using the general definition, we can say that an $\mathbb{R}^n$-valued random variable $X = (X_1, \ldots, X_n)$ is Gaussian if $\{X_1, \ldots, X_n\}$ is a Gaussian set.

7 Key Fact: uncorrelated implies independent for Gaussians

This is given in Problem 4.2.23 in the text: if (X, Y) is Gaussian and if X and Y are uncorrelated (meaning that $\Sigma_{X,Y} = 0$) then X and Y are independent. Note that the components of X need not be independent among themselves, and the components of Y need not be independent among themselves.