Chapter 3

Random Variables and Measurable Functions.

3.1 Measurability

Definition 42 *(Measurable function) Let* f *be a function from a measurable space* $(\Omega, \mathcal{F})$ *into the real numbers. We say that the function is* measurable *if for each Borel set* $B \in \mathcal{B}$ *, the set* $\{\omega; f(\omega) \in B\} \in \mathcal{F}$.

Definition 43 *(random variable) A* random variable X *is a measurable function from a probability space* $(\Omega, \mathcal{F}, \mathcal{P})$ *into the real numbers* $\Re$.

Definition 44 *(Indicator random variables) For an arbitrary set* $A \in \mathcal{F}$ *define* $I_A(\omega) = 1$ *if* $\omega \in A$ *and 0 otherwise. This is called an* indicator random variable.

Definition 45 *(Simple Random variables) Consider events* $A_i \in \mathcal{F}, i = 1, 2, 3, ..., n$ *such that* $\cup_{i=1}^n A_i = \Omega$. *Define* $X(\omega) = \sum_{i=1}^n c_i I_{A_i}(\omega)$ *where* $c_i \in \Re$. *Then* X *is measurable and is consequently a random variable. We normally assume that the sets* A_i *are disjoint. Because this is a random variable which can take only finitely many different values, then it is called* simple *and any random variable taking only finitely many possible values can be written in this form.*

Example 46 *(binomial tree) A stock, presently worth \$20, can increase each day by \$1 or decrease by \$1. We observe the process for a total of 5 days. Define* X *to be the value of the stock at the end of five days. Describe* $(\Omega, \mathcal{F})$ *and the function* $X(\omega)$ *. Define another random variable* Y *to be the value of the stock after 4 days.*

Define $X^{-1}(B) = \{\omega; X(\omega) \in B\}$. We will also sometimes denote this event $[X \in B]$. In the above example, define the events $X^{-1}(B)$ and $Y^{-1}(B)$ where $B = [20, \infty)$.

Then we have the following **properties.**
For any Borel sets $B_n \subset \Re$, and any random variable X,

1. $X^{-1}(\cup_n B_n) = \cup_n X^{-1}(B_n)$
2. $X^{-1}(\cap_n B_n) = \cap_n X^{-1}(B_n)$
3. $[X^{-1}(B)]^c = X^{-1}(B^c)$

These three properties together imply that for any class of sets $\mathcal{C}$, $X^{-1}(\sigma(\mathcal{C})) = \sigma(X^{-1}(\mathcal{C}))$. So X is measurable if, for all x, $\{\omega; X(\omega) \leq x\} \in \mathcal{F}$ (see Theorem 16 and Problem 3.16).

BEWARE: The fact that we use notation X^{-1} does not imply that the function X has an inverse in the usual sense. For example, if $X(\omega) = sin(\omega)$ for $\omega \in \Re$, then what is $X^{-1}([.5, 1])$?

Theorem 47 *(combining random variables) Suppose* $X_i, i = 1, 2, \ldots$ *are all (measurable) random variables. Then so are*

1. $X_1 + X_2 + X_3 + \ldots X_n$
2. X_1^2
3. cX_1 for any $c \in \mathcal{R}$
4. X_1X_2
5. $inf\ \{X_n; n \geq 1\}$
6. $\liminf X_n$
7. $sup\{X_n; n \geq 1\}$
8. $\limsup_{n\to\infty} X_n$

Proof. For 1. notice that $[X_1 + X_2 > x]$ if and only if there is a rational number q in the interval $X_1 > q > x - X_2$ so that $[X_1 > q]$ and $[X_2 > x - q]$. In other words

$[X_1+X_2 > x] = \cup_q [X_1 > q] \cap [X_2 > x-q]$ where the union is over all rational numbers q.

For 2, note that for $x \geq 0$,

$$[X_1^2 \leq x] = [X_1 \geq 0] \cap [X_1 \leq \sqrt{x}] \cup [X_1 < 0] \cap [X_1 \geq -\sqrt{x}].$$

For 3, in the case $c > 0$, notice that

$$[cX_1 \leq x] = [X_1 \leq \frac{x}{c}].$$

Finally 4 follows from properties 1, 2 and 3 since

$$X_1X_2 = \frac{1}{2}\{(X_1 + X_2)^2 - X_1^2 - X_2^2\}$$

For 5. note that $[\inf X_n \geq x] = \cap_{n=1}^{\infty}[X_n \geq x]$.

For 6. note that $[\liminf X_n \geq x] = [X_n > x - 1/m \quad \text{a.b.f.o.}]$ for all $m = 1, 2, ...$ so

$$[\liminf X_n \geq x] = \cap_{m=1}^{\infty} \liminf [X_n > x - 1/m].$$

The remaining two properties follow by replacing X_n by $-X_n$. ■

Definition 48 *(sigma-algebra generated by random variables) For X a random variable, define $\sigma(X) = \{X^{-1}(B); B \in \mathcal{B}\}$.*

$\sigma(X)$ is the smallest sigma algebra $\mathcal{F}$ such that X is a measurable function into $\Re$. The fact that it is a sigma-algebra follows from Theorem 16. Similarly, for a set of random variables $X_1, X_2, \ldots X_n$, the sigma algebra $\sigma(X_1, \ldots X_n)$ generated by these is the smallest sigma algebra such that all X_i are measurable.

Theorem 49 *$\sigma(X)$ is a sigma-algebra and is the same as $\sigma\{[X \leq x], \quad x \in \Re\}$.*

Definition 50 *A* Borel measurable function *f from $\Re \to \Re$ is a function such that $f^{-1}(B) \in \mathcal{B}$ for all $B \in \mathcal{B}$.*

For example if a function $f(x)$ is a continuous function from a subset of $\Re$ into a subset of $\Re$ then it is Borel measurable.

Theorem 51 *Suppose $f_i, i = 1, 2, \ldots$ are all Borel measurable functions. Then so are*

1. $f_1 + f_2 + f_3 + \ldots f_n$
2. f_1^2
3. cf_1 for any real number c.
4. $f_1 f_2$
5. $\inf\{f_n; n \geq 1\}$
6. $\liminf f_n$
7. $\sup\{f_n; n \geq 1\}$
8. $\lim_{n\to\infty} f_n$

Theorem 52 *If X and Y are both random variables, then Y can be written as a Borel measurable function of X, i.e. $Y = f(X)$ for some Borel measurable f if and only if*

$$\sigma(Y) \subset \sigma(X)$$

Proof. Suppose $Y = f(X)$. Then for an arbitrary Borel set B, $[Y \in B] = [f(X) \in B] = [X \in f^{-1}(B)] = [X \in B_2]$ for Borel set $B_2 \in \mathcal{B}$. This shows that $\sigma(Y) \subset \sigma(X)$.

For the converse, we assume that $\sigma(Y) \subset \sigma(X)$ and we wish to find a Borel measurable function f such that $Y = f(X)$. For fixed n consider the set $A_{m,n} = \{\omega; m2^{-n} \leq Y(\omega) < (m+1)2^{-n}\}$ for $m = 0, \pm 1, ...$ Since this set is in $\sigma(Y)$ it is also in $\sigma(X)$ and therefore can be written as $\{\omega; X(\omega) \in B_{m,n}\}$ for some Borel subset $B_{m,n}$ of the real line. Consider the function $f_n(x) = \sum_m m2^{-n} I(x \in B_{m,n})$. Clearly this function is defined so that $f_n(X)$ is close to Y, and indeed is within $\frac{1}{2^n}$ units of Y. The function we seek is obtained by taking the limit

$$f(x) = \lim_{n \to \infty} f_n(x).$$

We require two results, first that the limit exists and second that the limit satisfies the property $f(X) = Y$. Convergence of the sequence follows from the fact that for each x, the sequence $f_n(x)$ is monotonically increasing (this is Problem 22). The fact that $Y = f(X)$ follows easily since for each n, $f_n(X) \leq Y \leq f_n(X) + \frac{1}{2^n}$. Taking limits as $n \to \infty$ gives $f(X) \leq Y \leq f(X)$. ∎

Example 53 *Consider* $\Omega = [0,1]$ *with Lebesgue measure and define a random variable* $X(\omega) = a_1, a_2, a_3$ *(any three distinct real numbers) for* $\omega \in [0, 1/4], (1/4, 1/2], (1/2, 1]$ *respectively. Find* $\sigma(X)$. *Now consider a random variable* Y *such that* $Y(\omega) = 0$ *or* 1 *as* $\omega \in [0, 1/2], (1/2, 1]$ *respectively. Verify that* $\sigma(Y) \subset \sigma(X)$ *and that we can write* $Y = f(X)$ *for some Borel measurable function* $f(.)$.

3.2 Cumulative Distribution Functions

Definition 54 *The* cumulative distribution function *(c.d.f.) of a* random variable X *is defined to be the function* $F(x) = P[X \leq x]$, *for* $x \in \Re$. *Similarly, if* μ *is a measure on* $\Re$, *then the cumulative distribution function is defined to be* $F(x) = \mu(-\infty, x]$. *Note in the latter case, the function may take the value* ∞.

Theorem 55 *(Properties of the Cumulative Distribution Function)*

1. A c.d.f. $F(x)$ is non-decreasing. i.e. $F(y) \geq F(x)$ whenever $y \geq x$.
2. $F(x) \to 0$, as $x \to -\infty$.
3. When $F(x)$ is the c.d.f. of a random variable, $F(x) \to 1$, as $x \to \infty$.
4. $F(x)$ is right continuous. i.e. $F(x) = \lim F(x+h)$ as h decreases to 0.

Proof.

1. If $x \leq y$ then $X \leq x$ implies $X \leq y$ or in set theoretic terms $[X \leq x] \subset [X \leq y]$. Therefore $P(X \leq x) \leq P(X \leq y)$.

2. If X is a real-valued random variable then $[X = -\infty] = \varphi$ the empty set. Therefore for any sequence x_n **decreasing** to $-\infty$,

$$\begin{aligned}\lim F(x_n) &= \lim P(X \leq x_n) \\ &= P(\cap_{n=1}^{\infty}[X \leq x_n]) \text{ (since the sequence is nested)} \\ &= P(\varphi) = 0\end{aligned}$$

3. Again if X is a real-valued random variable then $[X < \infty] = \Omega$ and for any sequence x_n **increasing** to ∞,

$$\begin{aligned}\lim F(x_n) &= \lim P(X \leq x_n) \\ &= P(\cup_{n=1}^{\infty}[X \leq x_n]) \text{ (since the sequence is nested)} \\ &= P(\Omega) = 1.\end{aligned}$$

4. For any sequence h_n decreasing to 0,

$$\begin{aligned}\lim F(x+h_n) &= \lim P(X \leq x+h_n) \\ &= P(\cap_{n=1}^{\infty}[X \leq x+h_n]) \text{ (since the sequence is nested)} \\ &= P(X \leq x) = F(x)\end{aligned}$$

■

Theorem 56 *(existence of limits) Any bounded non-decreasing function has at most countably many discontinuities and possesses limits from both the right and the left. In particular this holds for cumulative distribution functions.*

Suppose we denote the limit of $F(x)$ from the left by $F(x-) = lim_h F(x-h)$ as h decreases to 0. Then $P[X < x] = F(x-)$ and $P[X = x] = F(x) - F(x-)$, the jump in the c.d.f. at the point x.

Definition 57 *Let x_i be any sequence of real numbers and p_i a sequence of non-negative numbers such that $\sum_i p_i = 1$. Define*

$$F(x) = \sum_{\{i; x_i \leq x\}} p_i. \tag{3.1}$$

This is the c.d.f. of a distribution which takes each value x_i with probability p_i. A discrete *distribution is one with for which there is a countable set S with $P[X \in S] = 1$. Any discrete distribution has cumulative distribution function of the form (3.1).*

Theorem 58 *If $F(x)$ satisfies properties 1-4 of Theorem 19, then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a random variable X defined on this probability space such that F is the c.d.f. of X.*

Proof. We define the probability space to be $\Omega = (0,1)$ with $\mathcal{F}$ the Borel sigma algebra of subsets of the unit interval and P the Borel measure. Define $X(\omega) = sup\{z; F(z) < \omega\}$. Notice that for any c, $X(\omega) > c$ implies $\omega > F(c)$. On the other hand if $\omega > F(c)$ then since F is right continuous, for some $\epsilon > 0$, $\omega > F(c+\epsilon)$ and this in turn implies that $X(\omega) \geq c + \epsilon > c$. It follows that $X(\omega) > c$ if and only if $\omega > F(c)$. Therefore $P[X(\omega) > c] = P[\omega > F(c)] = 1 - F(c)$ and so F is the cumulative distribution function of X. ∎

3.3 Problems

1. If $\Omega = [0,1]$ and P is Lebesgue measure, find $X^{-1}(C)$ where $C = [0, \frac{1}{2})$ and $X(\omega) = \omega^2$.
2. Define $\Omega = \{1,2,3,4\}$ and the sigma algebra $\mathcal{F} = \{\phi, \Omega, \{1\}, \{2,3,4\}\}$. Describe all random variables that are measurable on the probability space $(\Omega, \mathcal{F})$.
3. Let $\Omega = \{-2,-1,0,1,2\}$ and consider a random variable defined by $X(\omega) = \omega^2$. Find $\sigma(X)$,the sigma algebra generated by X.Repeat if $X(\omega) = |\omega|$ or if $X(\omega) = \omega + 1$.
4. Find two different random variables defined on the space $\Omega = [0,1]$ with Lebesgue measure which have exactly the same distribution. Can you arrange that these two random variables are independent of one another?
5. If X_i; $i = 1,2,...$ are random variables, prove that $max_{i \leq n} X_i$ is a random variables and that $limsup \frac{1}{n} \sum_i X_i$ is a random variable.
6. If X_i; $i = 1,2,...$ are random variables, prove that $X_1 X_2 ... X_n$ is a random variable.
7. Let Ω denote the set of all outcomes when tossing an unbiased coin three times. Describe the probability space and the random variable $X =$ *the number of heads observed*. Find the cumulative distribution function $P[X \leq x]$.
8. A number x is called a point of increase of a distribution function F if $F(x+\epsilon) - F(x-\epsilon) > 0$ for all $\epsilon > 0$. Construct a discrete distribution function such that every real number is a point of increase. (Hint: Can you define a discrete distribution supported on the set of all rational numbers?).
9. Consider a stock price process which goes up or down by a constant factor (e.g. $S_{t+1} = S_t u$ or $S_t d$ (where $u > 1$ and $d < 1$) with probabilities p and $1-p$ respectively (based on the outcome of the toss of a biased coin). Suppose we are interested in the path of the stock price from time $t = 0$ to time $t = 5$. What is a suitable probability space? What is $\sigma(S_3)$? What are the advantages of requiring that $d = 1/u$?

10. Using a Uniform random variable on the interval $[0,1]$, find a random variable X with distribution $F(x) = 1 - p^{\lfloor x \rfloor}$,$x > 0$, where $\lfloor x \rfloor$ denotes the floor or integer part of. Repeat with $F(x) = 1 - e^{-\lambda x}, x > 0, \lambda > 0$.

11. Suppose a coin with probability p of heads is tossed repeatedly. Let A_k be the event that a sequence of k or more consecutive heads occurs amongst tosses numbered $2^k, 2^k+1, \ldots, 2^{k+1}-1$. Prove that $P[A_k \;\; i.o.] = 1$ if $p \geq 1/2$ and otherwise it is 0.

    (*Hint:* Let E_i be the event that there are k consecutive heads beginning at toss numbered $2^k + (i-1)k$ and use the inclusion-exclusion formula.)

12. *The Hypergeometric Distribution* Suppose we have a collection (the *population*) of N objects which can be classified into two groups S or F where there are r of the former and $N-r$ of the latter. Suppose we take a random sample of n items without replacement from the population. Show the probability that we obtain exactly x S's is

$$f(x) = P[X = x] = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \;\; x = 0, 1, \ldots$$

    Show in addition that as $N \to \infty$ in such a way that $r/N \to p$ for some parameter $0 < p < 1$, this probability function approaches that of the *Binomial Distribution*

$$f(x) = P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \; 1, \ldots n$$

13. *The Negative Binomial distribution*

    The binomial distribution is generated by assuming that we repeated trials a fixed number n of times and then counted the total number of successes X in those n trials. Suppose we decide in advance that we wish a fixed number (k) of successes instead, and sample repeatedly until we obtain exactly this number. Then the number of trials X is random. Show that the probability function is:

$$f(x) = P[X = x] = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, \ldots$$

14. Let $g(u)$ be a cumulative distribution function on $[0,1]$ and $F(x)$ be the cumulative distribution function of a random variable X. Show that we can define a *deformed* cumulative distribution function such that $G(x) = g(F(x))$ at at all continuity points of $g(F(x))$. Describe the effect of this transformation when

$$g(u) = \Phi\left(\Phi^{-1}(u) - \alpha\right)$$

    for Φ the standard normal cumulative distribution function. Take a special case in which F corresponds to the $N(2,1)$ cumulative distribution function.

15. Show that if X has a continuous c.d.f. $F(x)$ then the random variable $F(X)$ has a uniform distribution on the interval $[0,1]$.

16. Show that if $\mathcal{C}$ is a class of sets which generates the Borel sigma algebra in $\mathcal{R}$ and X is a random variable then $\sigma(X)$ is generated by the class of sets
$$\{X^{-1}(A); A \in \mathcal{C}\}.$$

17. Suppose that $X_1, X_2, ...$are independent Normal(0,1) random variables and $S_n = X_1 + X_2 + ... + X_n$. Use the Borel Cantelli Lemma to prove the strong law of large numbers for normal random variables. i.e. prove that for and $\varepsilon > 0$,
$$P[S_n > n\varepsilon \quad i.o.] = 0.$$
Note: you may use the fact that if $\Phi(x)$ and $\phi(x)$ denote the standard normal cumulative distribution function and probability density function respectively, $1 - \Phi(x) \leq Cx\phi(x)$ for some constant C. Is it true that $P[S_n > \sqrt{n}\varepsilon \quad i.o.] = 0$?

18. Show that the following are equivalent:

    (a) $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for all x, y

    (b) $P(X \in A, Y \in B) = P(X \in A)P(X \in B)$ for all Borel subsets of the real numbers A, B.

19. Let X and Y be independent random variables. Show that for any Borel measurable functions f, g on $\mathcal{R}$, the random variables $f(X)$ and $g(Y)$ are independent.

20. Show that if A is an uncountable set of non-negative real numbers, then there is a sequence of elements of A, $a_1, a_2, ...$ such that $\sum_{i=1}^{\infty} a_i = \infty$.

21. Mrs Jones made a rhubarb crumble pie. While she is away doing heart bypass surgery on the King of Tonga, her son William (graduate student in Stat-Finance) comes home and eats a random fraction X of the pie. Subsequently her daughter Wilhelmina (PhD student in Stat-Bio) returns and eats a random fraction Y of the remainder. When she comes home, she notices that more than half of the pie is gone. If one person eats more than a half of a rhubard-crumble pie, the results are a digestive catastrophe. What is the probability of such a catastrophe if X and Y are independent uniform on $[0,1]$?

22. Suppose for random variables Y and X, $\sigma(Y) \subset \sigma(X)$. Define sets by
$$A_{m,n} = \{\omega; m2^{-n} \leq Y(\omega) < (m+1)2^{-n} \text{ for } m = 0, \pm 1, ...$$
and define a function f_n by
$$f_n(x) = \sum_m m2^{-n} I(x \in B_{m,n})$$

where

$$[X \in B_{m,n}] = A_{m,n}.$$

Prove that the sequence of functions f_n is non-decreasing in n.

23. Let (Ω, F, P) be the unit interval $[0,1]$ together with the Borel subsets and Borel measure. Give an example of a function from $[0,1]$ into $\mathcal{R}$ which is NOT a random variable.

24. Let (Ω, F, P) be the unit interval $[0,1]$ together with the Borel subsets and Borel measure. Let $0 \leq a < c < c < d \leq 1$ be arbitrary real numbers. Give and example of a sequence of events $A_n, n = 1, 2, ...$ such that the following all hold:

$$P(\liminf A_n) = a$$
$$\liminf P(A_n) = b$$
$$\limsup P(A_n) = c$$
$$P(\limsup A_n) = d$$

25. Let $A_n, n = 1, 2, ...$ be a sequence of events such that A_i and A_j are independent whenever

$$|i - j| \geq 2$$

and $\sum_n P(A_n) = \infty$. Prove that

$$P(\limsup A_n) = 1$$

26. For each of the functions below find the smallest sigma-algebra for which the function is a random variable. $\Omega = \{-2, -2, 0, 1, 2\}$ and

   (a) $X(\omega) = \omega^2$

   (b) $X(\omega) = \omega + 1$

   (c) $X(\omega) = |\omega|$

27. Let $\Omega = [0,1]$ with the sigma-algebra $\mathcal{F}$ of Borel subsets B contained in this unit interval which have the property that $B = 1 - B$.

   (a) Is $X(\omega) = \omega$ a random variable with respect to this sigma-algebra?

   (b) Is $X(\omega) = |\omega - \frac{1}{2}|$ a random variable with respect to this sigma-algebra?

28. Suppose Ω is the unit square in two dimensions together with Lebesgue measure and for each $\omega \in \Omega$, we define a random variable $X(\omega)$ = minimum distance to an edge of the square. Find the cumulative distribution function of X and its derivative.

29. Suppose that X and Y are two random variables on the same probability space with joint distribution

$$P(X = m, Y = n) = \begin{cases} \frac{1}{2^{m+1}} & \text{if } m \geq n \\ 0 & \text{if } m < n \end{cases}.$$

Find the marginal cumulative distribution functions of X and Y.