

A collection  $\mathcal{F}$  of subsets of  $\Omega$  is a field (i) if  $A, B \in \mathcal{F}$  then  $A \cup B \in \mathcal{F}$  and  $A \cap B \in \mathcal{F}$ , (ii) if  $A \in \mathcal{F}$  then  $A^c \in \mathcal{F}$ , (iii)  $\emptyset \in \mathcal{F}$ .  
 $\sigma$ -field: (i)  $\emptyset \in \mathcal{F}$ , (ii) if  $A_1, A_2, \dots \in \mathcal{F}$  then  $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ , (iii) if  $A \in \mathcal{F}$  then  $A^c \in \mathcal{F}$ .

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad (A \cup B)^c = A^c \cap B^c \quad (A \cap B)^c = A^c \cup B^c \quad A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B)$$

$$A \cap B^c = A - (A \cap B)$$

$$P(\emptyset) = 0 \quad P(\Omega) = 1 \quad P(A^c) = 1 - P(A) \quad \text{If } A \subseteq B \text{ then } P(A) \leq P(B).$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad P(A \cap B^c) = P(A) - P(A \cap B) = P(A - B)$$

$$\text{If } A_i \text{ are disjoint then } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \text{ In general } P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i).$$

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad P(A) = P(A | B)P(B) + P(A | B^c)P(B^c) \quad P(C | A \cap B)P(A \cap B) = P(C | A \cap B)P(B | A)P(A)$$

$$P(ABC) = P(A \cap B \cap C) = \frac{P(A \cap B \cap C)}{P(A \cap B)}P(A \cap B) = P(C | A \cap B)P(A \cap B)$$

$$P(C) = P(C | A \cap B)P(A \cap B) + P(C | A \cap B^c)P(A \cap B^c) + P(C | A^c \cap B)P(A^c \cap B) + P(C | A^c \cap B^c)P(A^c \cap B^c)$$

$$\text{Bayes' formula } P(A | X) = \frac{P(X | A)P(A)}{P(X | A)P(A) + P(X | B)P(B) + P(X | C)P(C)}$$

$$\text{Independence } P(A \cap B) = P(A)P(B) \quad P(A | B) = P(A)$$

$$\text{Conditional independence } P(A \cap B | C) = P(A | C)P(B | C)$$

$$\text{"lim inf and lim sup are events."} \quad \liminf A_k = \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} A_j \quad \limsup A_k = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$$

$$\bigcap_{k=1}^{\infty} A_k = \{\omega : \omega \in A_k \text{ for all } k\}$$

$$\liminf A_k = \{\omega : \omega \in A_k \text{ for all } k > n \text{ for some } n\}$$

$$\limsup A_k = \{\omega : \omega \in A_k \text{ for infinitely many } k\}$$

$$\bigcup_{k=1}^{\infty} A_k = \{\omega : \omega \in A_k \text{ for at least one } k\}$$

$$\bigcap_{k=1}^{\infty} A_k \subseteq \liminf A_k \subseteq \limsup A_k \subseteq \bigcup_{k=1}^{\infty} A_k$$

$$\liminf(A_n \cap B_n) = (\liminf A_n) \cap (\liminf B_n) \quad (\limsup A_k)^c = \liminf(A_k^c)$$

$$\limsup(A_n \cup B_n) = (\limsup A_n) \cup (\limsup B_n) \quad \liminf A_n \subseteq \liminf(A_n \cup B_n)$$

Borel-Cantelli lemma: Let  $A_k \in \mathcal{F}$  such that  $\sum_{k=1}^{\infty} P(A_k) < \infty$ . Then  $P(\limsup A_k) = 0$ .

i.o. "infinitely often" is the same as lim sup

$$\text{Increasing sequence of events: } A_1 \subseteq A_2 \subseteq \dots \text{ then } \lim_{i \rightarrow \infty} A_i = \bigcup_{i=1}^{\infty} A_i$$

$$\text{Decreasing sequence of events: } A_1 \supseteq A_2 \supseteq \dots \text{ then } \lim_{i \rightarrow \infty} A_i = \bigcap_{i=1}^{\infty} A_i$$

$$\text{Continuity theorem: } \lim_{k \rightarrow \infty} P(A_k) = P(\lim_{k \rightarrow \infty} A_k)$$

$$I_A = 1 - I_{A^c} \quad I_{A \cap B} = I_A \cdot I_B \quad I_{A \cup B} = \max(I_A, I_B) = I_A + I_B - I_A I_B \quad I_A^2 = I_A$$

If  $F$  is a distribution function then (i)  $\lim_{x \rightarrow -\infty} F(x) = 0$ ,  $\lim_{x \rightarrow \infty} F(x) = 1$ , (ii) if  $x < y$  then  $F(x) \leq F(y)$ , (iii)  $F(x)$  is right continuous,  $\lim_{h \downarrow 0} F(x+h) = F(x)$ . Converse is true, any function that satisfies (i), (ii), and (iii) is a distribution function of some random variable. If  $F(x) \leq F(y)$  then  $\{\omega : X(\omega) \leq x\} \subseteq \{\omega : Y(\omega) \leq y\}$ .

$X$  is discrete if it can have only countably many values.

If  $f(x)$  is a probability mass function then (i)  $0 \leq f(x) \leq 1$  for all  $x$ , (ii)  $\sum_x f(x) = 1$ . Converse is true, any function that satisfies (i) and (ii) is a p.m.f. of some random variable.

$$f(x) = P(X = x) = F(x) - \lim_{h \downarrow 0} F(x-h) \quad F(x) = P(X \leq x) = \sum_{y \leq x} f(y)$$

$X$  is a continuous random variable if there is an  $f(x)$  such that (i)  $f(x) \geq 0$  for all  $x$ , (ii)  $F(x) = \int_{-\infty}^x f(y) dy$ . The function  $f(x)$  is called the density of  $X$ . Density of a random variable has to satisfy two conditions (i)  $f(x) \geq 0$ , (ii)  $\int_{-\infty}^{\infty} f(x) = 1$ .

Remark  $f(x) = dF/dx$ . "A density is not a probability like mass is."

$$P(X \leq x) = F(x) \quad P(X > x) = 1 - F(x) \quad P(X < x) = \lim_{h \downarrow 0} F(x - h) \quad P(X \geq x) = 1 - \lim_{h \downarrow 0} F(x - h)$$

$$P(a < X < b) = \lim_{h \downarrow 0} F(b - h) - F(a) \quad P(a \leq X \leq b) = F(b) - \lim_{h \downarrow 0} F(a - h) \quad P(X = x) = F(x) - \lim_{h \downarrow 0} F(x - h)$$

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = 1 \quad \sum_{k=1}^{\infty} \frac{1}{k} = \infty \quad \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \quad \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90} \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{k=0}^n p^k = \frac{1-p^{n+1}}{1-p} \quad \sum_{k=0}^{\infty} p^k = \frac{1}{1-p} \quad \sum_{k=1}^{\infty} p^k = \frac{p}{1-p} \quad \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$$

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \left(\frac{k}{n}\right)^m = \int_0^1 x^m dx = \frac{1}{m+1} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$\log ab = \log a + \log b \quad \log(a/b) = \log a - \log b \quad \log a^b = b \log a \quad \log_2 a = \frac{\log a}{\log 2} \quad \lim_{x \rightarrow -\infty} \tan^{-1} x = -\frac{\pi}{2}$$

$$\int e^{ax} = \frac{e^{ax}}{a} \quad \int x e^{ax} = \frac{e^{ax}(ax-1)}{a^2} \quad \int \frac{du}{a^2+u^2} = \frac{1}{a} \tan^{-1} \frac{u}{a} \quad \int \frac{du}{u} = \log |u|$$

$$\text{Expected value } EX = \sum x f(x) \quad Eg(x) = \sum g(x) f(x) \quad E(aX+b) = aEX+b$$

Theorem (i)  $E1 = 1$ , (ii)  $E(aX+bY) = aEX+bEY$ , (iii) if  $X \geq 0$  then  $EX \geq 0$ .

$$\text{Var } X = E((X-EX)^2) \quad \text{Var } X = E(X^2) - (EX)^2 \quad \text{Var}(aX+b) = a^2 \text{Var } X$$

If  $X$  and  $Y$  are independent then  $E(XY) = E(X)E(Y)$ . Converse is not true in general.

$$\text{Cov}(X, Y) = E[(X-EX)(Y-EY)] = E(XY) - E(X)E(Y) \quad \text{If } \text{Cov}(X, Y) = 0 \text{ then } X \text{ and } Y \text{ are uncorrelated.}$$

If  $X$  and  $Y$  are uncorrelated then  $\text{Var}(X+Y) = \text{Var } X + \text{Var } Y$

$$\text{Correlation coefficient } \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

$$E(X(X-1)) = E(X^2) - EX \quad \text{Var } X = E(X^2) - (EX)^2 = E(X(X+1)) - EX - (EX)^2$$

Bernoulli( $p$ ) —  $f(1)$  is the probability of heads,  $f(0)$  is the probability of tails.

Binomial( $n, p$ ) —  $f(k)$  is the probability of  $k$  heads in  $n$  tosses.

Poisson( $\lambda$ )

Geometric( $p$ ) —  $f(k)$  is the probability that it takes  $k$  tosses to get a head.

Negative Binomial( $n, p$ ) —  $f(k)$  is the probability of  $k$  tosses to get  $n$  heads.

Hypergeometric( $N, b, n$ ) — Urn with  $N$  balls,  $b$  are black and  $N-b$  are white. Draw  $n$  balls without replacement.  $f(k)$  is the probability of  $k$  black balls.

Joint distribution function  $F(x, y) = P(X \leq x \text{ and } Y \leq y)$  Joint mass function  $f(x, y) = P(X = x \text{ and } Y = y)$

Joined distribution function properties (i)  $\lim_{x, y \rightarrow -\infty} F(x, y) = 0$ ,  $\lim_{x, y \rightarrow \infty} F(x, y) = 1$ , (ii) if  $(x_1, y_1) \leq (x_2, y_2)$  then  $F(x_1, y_1) \leq F(x_2, y_2)$ , (iii) continuous from above,  $\lim_{u, v \downarrow 0} F(x+u, y+v) = F(x, y)$ . Property (ii) means  $F(x, y)$  is nondecreasing. Another way of putting it is, for any  $a_1 < a_2$ ,  $b_1 < b_2$ , it must be true that  $F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1) \geq 0$ . Any function that satisfies (i), (ii), and (iii) is a joined distribution function.

$$\lim_{y \rightarrow \infty} F(x, y) = F_X(x) = P(X \leq x) \quad \lim_{x \rightarrow \infty} F(x, y) = F_Y(y) = P(Y \leq y)$$

For random vectors,  $x \leq y$  means that  $x_1 \leq y_1$  and  $x_2 \leq y_2$ .

$$F(x, y) \text{ is jointly continuous if } F(x, y) = \int_{u=-\infty}^x \int_{v=-\infty}^y f(u, v) du dv$$

Discrete random variables  $X$  and  $Y$  are independent iff  $f(x, y)$  can be factored, that is,  $f(x, y) = g(x)h(y)$ .

For any intervals  $A$  and  $B$ , if  $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$  then  $A$  and  $B$  are independent.

Transformation theorem: If  $X$  and  $Y$  are independent, then  $g(X)$  and  $h(Y)$  are also independent.

$$P(X \in A) = \sum_{x \in A} f(x)$$

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1)$$

“We know that for a decreasing sequence of sets, the limit exists and it’s the empty set.”