Probability and Random Processes - 3rd ed.
Grimmett and Stirzaker (2001)

Robin Aldridge-Sutton

July 1, 2021

1 Events and their probabilities

- Sample space - A set Ω of all possible outcomes of an experiment
- σ-field - A collection $\mathcal{F}$ of subsets of Ω with
  - $\emptyset \in \mathcal{F}$
  - $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  - $A_1, A_2, ... \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$
- Probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ - A function $\mathbb{P} : \mathcal{F} \to [0, 1]$ with
  - $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
  - $A_1, A_2, ... \in \mathcal{F}$ and $\forall i \neq j, A_i \cap A_j = \emptyset \Rightarrow \mathbb{P}\big(\bigcup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$
- Probability space - A triple $(\Omega, \mathcal{F}, \mathbb{P})$ comprising a set Ω, a σ-field $\mathcal{F}$, and a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$
- Conditional probability that event A occurs given that event B occurs
  - $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \mathbb{P}(B) > 0$
- Lemma - $B_1, B_2, ...$ a partition of Ω with $\mathbb{P}(B_i) > 0$ for all i $\Rightarrow$
  - $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$
- Independent events $\{A_i : i \in I\}$
  - $\mathbb{P}(\cap_{i \in J} A_i) = \prod_{i \in J} \mathbb{P}(A_i)$ for all finite $J \subset I$

2 Random variables and their distributions

- Random variable ($\mathcal{F}$-measurable function) - A function $X : \Omega \to \mathbb{R}$ with
  - $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for each $x \in \mathcal{R}$
- Distribution function of a random variable X - A function $F : \mathbb{R} \to [0,1]$ given by
  - $F(x) = \mathbb{P}(X \leq x)$
- Discrete random variable - A random variable X taking values in some countable subset $\{x_1, x_2, ...\}$ of $\mathbb{R}$ and has (probability) mass function $f : \mathbb{R} \to [0,1]$ given by
  - $f(x) = \mathbb{P}(X = x)$
- Continuous random variable - A random variable X with distribution function expressible as
  - $F(x) = \int_{-\infty}^{x} f(u)du, x \in \mathbb{R}$
  - $f : \mathbb{R} \to [0, \infty)$ is called the (probability) density function of X
- Joint distribution function of a random vector $\mathbf{X} = \{X_1, X_2, ..., X_n\}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the function $F_{\mathbf{X}} : \mathbb{R} \to [0,1]$ given by
  - $F_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^n$
  - $\mathbf{x} \leq \mathbf{y} \iff x_i \leq y_i$ for $i \in \{1, ..., n\}$

3 Discrete random variables

- Independent discrete random variables - Events $\{X = x\}$ and $\{Y = y\}$ are independent
- Expected value of discrete random variable X with mass function f
  - $\mathbb{E}(X) = \sum_{x:f(x)>0} xf(x)$,
  - where this sum is absolutely convergent so that the order of the terms is irrelevant
- Lemma - If X has mass function f, and $g : \mathcal{R} \to \mathcal{R}$ then
  - $\mathbb{E}(g(X)) = \sum_x g(x)f(x)$,
  - where this sum is absolutely convergent
- k th moment m_k of X
  - $m_k = \mathbb{E}(X^k)$
  - $\sigma_k = \mathbb{E}((X - m_1)^k)$ (kth central moment)

- Theorem - The expectation operator $\mathbb{E}$ has the property
  - $a, b \in \mathbb{R} \Rightarrow \mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$
- Joint distribution function $F : \mathbb{R}^2 \to [0,1]$, and mass function $f : \mathbb{R}^2 \to [0,1]$, of discrete random variables X and Y
  - $F(x, y) = \mathbb{P}(X \leq x \text{ and } Y \leq y)$
  - $f(x, y) = \mathbb{P}(X = x \text{ and } Y = y)$
- Lemma - Discrete random variables X and Y are independent if and only if
  - $f_{X,Y}(x, y) = f_X(x)f_Y(y)$, or
  - $f_{X,Y}(x, y) = g(x)h(y)$, for all $x, y \in \mathbb{R}$
- Conditional distribution function $F_{Y|X}(.|x)$, and conditional (probability) mass function $f_{Y|X}(.|x)$, of Y given $X = x$
  - $F_{Y|X}(y|x) = \mathbb{P}(Y \leq y|X = x)$
  - $f_{Y|X}(y|x) = \mathbb{P}(Y = y|X = x)$,
  - $\mathbb{P}(X = x) > 0$
- Conditional expectation of Y given X
  - $\mathbb{E}(Y|X) = \sum_y y f_{Y|X}(y|X)$
  - $\mathbb{E}_X(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$
- Sum of random variables - X and Y independent $\Rightarrow$
  - $\mathbb{P}(X + Y = z) = f_{X+Y}(z) = (f_X * f_Y)(z)$
  - $= \sum_x f_X(x)f_Y(z - x) = \sum_y f_X(z - y)f_Y(y)$
  - $f_X * f_Y$ is the convolution of the mass functions of X and Y

4 Continuous random variables

- Independent continuous random variables X and Y
  - $\{X \leq x\}$ and $\{Y \leq y\}$ are independent events for all $x, y \in \mathbb{R}$
- Expectation of a continuous random variable X with density function f
  - $\mathbb{E}X = \int_{-\infty}^{\infty} x f(x) dx$
  - whenever this integral exists
- Theorem - X and $g(X)$ continuous random variables
  - $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x) dx$

- Joint distribution function $F : \mathbb{R}^2 \to [0, 1]$, and density function $f : \mathbb{R}^2 \to [0, \infty)$, of continuous random variables X and Y
  - $F(x, y) = \mathbb{P}(X \leq x, Y \leq y)$
  - $F(x, y) = \int_{v=-\infty}^{y} \int_{u=-\infty}^{x} f(u, v) du dv$
  - for each $x, y \in \mathbb{R}$
- Conditional distribution function $F_{Y|X}(.|x)$, and conditional (probability) density function $f_{Y|X}(.|x)$, of continuous random variable Y given $X = x$
  - $F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)} dv$,
  - $\mathbb{P}(Y|X = x)$ does not exist because $\mathbb{P}(X = x) = 0$,
  - $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$,
  - $f_X(x) > 0$
- Conditional expectation of Y given X
  - $\mathbb{E}(Y|X) = \int_{-\infty}^{\infty} y f_{Y|X}(y|X) dy$,
  - $\mathbb{E}_X(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$
- Theorem (Change of variables)
  - $g : \mathbb{R}^n \to \mathbb{R}$,
  - $T : (x_1, x_2) \in A \subseteq D \subseteq \mathbb{R}^2 \to (y_1, y_2) \in B \subseteq R \subseteq \mathbb{R}^2$ one-to-one, and
  - $J(y_1, y_2)$ the determinant of the matrix of first-order partial derivatives of its inverse (both matrix and determinant called the Jacobian) $\Rightarrow$
  - $\int \int_A g(x_1 x_2) dx_1 dx_2 = \int \int_B g(x_1(y_1, y_2), x_2(y_1, y_2)) |J(y_1, y_2)| dy_1 dy_2$
- Corrollary - X_1, X_2 joint density function f and $(Y_1, Y_2) = T(X_1, X_2) \Rightarrow$
  - $f_{Y_1,Y_2}(y_1, y_2) = f(x_1(y_1, y_2), x_2(y_1, y_2)) |J(y_1, y_2)|$,
  - if $(y_1, y_2) \in R$, and 0 otherwise.
- Sum of random variables - X and Y independent $\Rightarrow$
  - $f_{X+Y}(z) = (f_X * f_Y)(z)$
  - $= \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy$
  - $f_X * f_Y$ is the convolution of the mass functions of X and Y

5 Generating functions and their applications

- Generating function of a sequence a - The function (power series) G_a defined by
  - $G_a(s) = \sum_{i=0}^{\infty} a_i s^i$,
  - for $s \in \mathbb{R}$ for which the sum converges
  - The sequence a is recoverable by $a_i = \frac{G_a^{(i)}(0)}{i!}$
- (Probability) Generating function - The generating function $G(s)$, of the probability mass function $\{f(i) : i \geq 0\}$, of a discrete random variable X, taking values in $\{0, 1, ...\}$, taking the form
  - $G(s) = \mathbb{E}(s^X) = \sum_i s^i f(i)$
- Convlution of real sequences $\{a_i : i \geq 0\}$ and $\{b_i : i \geq 0\}$ - The sequence $\{c_i : i \geq 0\}$ defined by
  - $c_n = a_0 b_n + ... + a_n b_0$
  - $c = a * b$
- The generating function of $c = a * b$ is
  - $G_c(s) = \sum_{n=0}^{\infty} c_n s^n = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) s^n$
  - $= \sum_{i=0}^{\infty} a_i s^i \sum_{i=0}^{\infty} b_{n-i} s^{n-i}$
  - $G_c(s) = G_a(s) G_b(s)$
- Theorem - X and Y independent $Rightarrow$
  - $G_{X+Y}(s) = G_X(s) G_Y(s)$

6 Random processes (Ch 8)

- Random process - A family $X = \{X_t : t \in T\}$ of random variables indexed by an index set T, which map a sample space Ω into a state space S
- Sample path (Realization) of a random process X at a fixed $\omega \in \Omega$ - The collection $\{X_t(\omega) : t \in T\}$ of members of S
- Finite-dimensional distributions of a random process X - The collection $\{F_{\mathbf{t}}\}$ of joint distribution functions $F_{\mathbf{t}} : \mathbb{R}^n \to [0, 1]$ given by $F_{\mathbf{t}}(\mathbf{x}) = \mathbb{P}(X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n)$ as $\mathbf{t}$ ranges over all vectors $(t_1, t_2, ..., t_n)$ of members of T
- Strongly stationary process - A process $X = \{X_t : t \geq 0\}$, taking values in $\mathbb{R}$ with families $\{X(t_1), ..., X(t_n)\}$ and $\{X(t_1 + h), ..., X(t_n + h)\}$ having the same distribution for all $t_1, ..., t_n$ and $h > 0$

- Weakly stationary process - A process $X = \{X_t : t \geq 0\}$ with
  - $\mathbb{E}(X(t_1)) = \mathbb{E}(X(t_1))$
  - $cov(X(t_1), X(t_2)) = cov(X(t_1 + h), X(t_2 + h))$
  - for all t_1, t_2, and $h > 0$

7 Diffusion processes (Ch 13)

- Wiener process - A real-valued Gaussian process $W = \{W(t) : t \geq 0\}$, starting from $W(0) = w$, with
  - Independent increments
  - W(s + t) - $W(s) \sim N(0, \sigma^2 t)$ for all $s, t \geq 0$ where σ^2 is a positive constant
  - The sample paths of W are continuous