Todennäköisyys ja statistinen fysiikka

Probability and statistical physics

Antti Kemppainen <Antti.H.Kemppainen@helsinki.fi> B416, ma 14-15
Kalle Kytölä <Kalle.Kytola@helsinki.fi> A406, to 16-17

Helsingin yliopisto, Matematiikan ja tilastotieteen laitos

kevätlukukausi 2014
periodit III & IV

Kolmogorov's axioms: probability space

- a set Ω ("sample space", "set of possible outcomes")
- σ-algebra $\mathcal{F}$: collection of subsets of Ω ("events")

  (σ-1) $\varnothing \in \mathcal{F}$

  (σ-2) $E \in \mathcal{F} \;\Rightarrow\; \Omega \setminus E \in \mathcal{F}$

  (σ-2) $E_n \in \mathcal{F}\ (n \in \mathbb{N}) \;\Rightarrow\; \bigcup_{n\in\mathbb{N}} E_n \in \mathcal{F}$
- probability measure $\mathrm{P}\colon \mathcal{F} \to [0,1]$

  (P-1) $\mathrm{P}[\Omega] = 1$

  (P-2) $E_n \in \mathcal{F}\ (n \in \mathbb{N})$ disjoint, i.e. $E_n \cap E_m = \varnothing$ for $n \neq m \;\Rightarrow$ $\mathrm{P}[\bigcup_{n\in\mathbb{N}} E_n] = \sum_{n\in\mathbb{N}} \mathrm{P}[E_n]$

	probability theory, $(\Omega, \mathcal{F}, \mathrm{P})$
	event, $E \in \mathcal{F}$ $(E \subset \Omega)$
	probability, $\mathrm{P}[E]$
	random variable, $X\colon \Omega \to \mathbb{R}$
	expected value, $\mathrm{E}[X]$
	a.s. = almost surely

Kolmogorov's axioms: probability space

- a set Ω ("sample space", "set of possible outcomes") $\mathfrak{X}$
- σ-algebra $\mathcal{F}$: collection of subsets of Ω ("events") Σ

  (σ-1) $\varnothing \in \mathcal{F}$

  (σ-2) $E \in \mathcal{F} \;\Rightarrow\; \Omega \setminus E \in \mathcal{F}$

  (σ-2) $E_n \in \mathcal{F}$ $(n \in \mathbb{N})$ $\Rightarrow$ $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{F}$
- probability measure $\mathrm{P}\colon \mathcal{F} \to [0,1]$ $\mu\colon \Sigma \to [0,+\infty]$

  (P-1) $\mathrm{P}[\Omega] = 1$ $\mu[\varnothing] = 0$

  (P-2) $E_n \in \mathcal{F}$ $(n \in \mathbb{N})$ disjoint, i.e. $E_n \cap E_m = \varnothing$ for $n \neq m$ $\Rightarrow$
  $\mathrm{P}[\bigcup_{n\in\mathbb{N}} E_n] = \sum_{n\in\mathbb{N}} \mathrm{P}[E_n]$

measure theory, $(\mathfrak{X}, \Sigma, \mu)$	probability theory, $(\Omega, \mathcal{F}, \mathrm{P})$
measurable set, $A \in \Sigma$ $(A \subset \mathfrak{X})$	event, $E \in \mathcal{F}$ $(E \subset \Omega)$
measure, $\mu[A]$	probability, $\mathrm{P}[E]$
measurable function, $f\colon \mathfrak{X} \to \mathbb{R}$	random variable, $X\colon \Omega \to \mathbb{R}$
integral, $\int_{\mathfrak{X}} f \,\mathrm{d}\mu$	expected value, $\mathrm{E}[X]$
a.e. = almost everywhere	a.s. = almost surely

Expected value: examples/construction

- Indicator $\mathbb{I}_E$ of an event E

$$\mathbb{I}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}, \qquad \mathrm{E}[\mathbb{I}_E] = \mathrm{P}[E]$$

Expected value: examples/construction

- Indicator $\mathbb{I}_E$ of an event E

$$\mathbb{I}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}, \qquad \mathrm{E}[\mathbb{I}_E] = \mathrm{P}[E]$$

- Simple random variable X (finitely many possible values)

$$X = \sum_{j=1}^{n} x_j \mathbb{I}_{E_j} \qquad (E_i \cap E_j = \varnothing \ \forall i \neq j), \qquad \mathrm{E}[X] = \sum_{j=1}^{n} x_j \mathrm{P}[E_j]$$

Expected value: examples/construction

- Indicator $\mathbb{I}_E$ of an event E

$$\mathbb{I}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}, \qquad \mathrm{E}[\mathbb{I}_E] = \mathrm{P}[E]$$

- Simple random variable X (finitely many possible values)

$$X = \sum_{j=1}^{n} x_j \mathbb{I}_{E_j} \qquad (E_i \cap E_j = \varnothing \ \forall i \neq j), \qquad \mathrm{E}[X] = \sum_{j=1}^{n} x_j \mathrm{P}[E_j]$$

- Non-negative random variable $X \geq 0$
  approximate from below by $(X_n)_{n \in \mathbb{N}}$ simple, $X_n \nearrow X$

$$\text{e.g. } X_n = \sum_{j=1}^{4^n - 1} \frac{j}{2^n} \mathbb{I}_{\{j 2^{-n} \leq X < (j+1) 2^{-n}\}} + 2^n \mathbb{I}_{\{X \geq 2^n\}}, \quad \mathrm{E}[X] = \lim_{n \to \infty} \mathrm{E}[X_n]$$

Expected value: examples/construction

- Indicator $\mathbb{I}_E$ of an event E

$$\mathbb{I}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}, \qquad \mathrm{E}[\mathbb{I}_E] = \mathrm{P}[E]$$

- Simple random variable X (finitely many possible values)

$$X = \sum_{j=1}^{n} x_j \mathbb{I}_{E_j} \qquad (E_i \cap E_j = \varnothing \ \forall i \neq j), \qquad \mathrm{E}[X] = \sum_{j=1}^{n} x_j \mathrm{P}[E_j]$$

- Non-negative random variable $X \geq 0$
  approximate from below by $(X_n)_{n\in\mathbb{N}}$ simple, $X_n \nearrow X$

$$\text{e.g. } X_n = \sum_{j=1}^{4^n-1} \frac{j}{2^n} \mathbb{I}_{\{j2^{-n} \leq X < (j+1)2^{-n}\}} + 2^n \mathbb{I}_{\{X \geq 2^n\}}, \qquad \mathrm{E}[X] = \lim_{n\to\infty} \mathrm{E}[X_n]$$

- Integrable random variable $X \in L^1(\mathrm{P})$, i.e., $\mathrm{E}[|X|] < \infty$

$$X = X_+ - X_- \text{ with } X_+, X_- \geq 0, \qquad \mathrm{E}[X] = \mathrm{E}[X_+] - \mathrm{E}[X_-]$$

Useful results from measure theory

Exchange of limit and integration:

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n\big)\,\mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n\,\mathrm{d}\mu\Big)$$

Useful results from measure theory

Exchange of limit and integration: $(f_n)_{n\in\mathbb{N}}$ m$\underline{\text{able}}$ f$\underline{\text{tions}}$ on $\mathfrak{X}$

- Monotone convergence theorem

  f_n non-negative, and $f_n \uparrow f$

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n\big)\,\mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n \,\mathrm{d}\mu \Big)$$

Useful results from measure theory

Exchange of limit and integration: $(f_n)_{n\in\mathbb{N}}$ m$\underline{\text{able}}$ f$\underline{\text{tions}}$ on $\mathfrak{X}$

- Monotone convergence theorem

  f_n non-negative, and $f_n \uparrow f$

- Dominated convergence theorem

  $|f_n| \leq g$ for some $g \in L^1(\mu)$, and $f_n \to f$

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n\big)\,\mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n\,\mathrm{d}\mu\Big)$$

Useful results from measure theory

Exchange of limit and integration: $(f_n)_{n\in\mathbb{N}}$ m$\underline{\text{able}}$ f$\underline{\text{tions}}$ on $\mathfrak{X}$

- Monotone convergence theorem

  f_n non-negative, and $f_n \uparrow f$

- Dominated convergence theorem

  $|f_n| \leq g$ for some $g \in L^1(\mu)$, and $f_n \to f$

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n \big) \, \mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n \, \mathrm{d}\mu \Big)$$

Exchange of order of integrations:

$$\Rightarrow \quad \int_{\mathfrak{X}_1 \times \mathfrak{X}_2} f \, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{\mathfrak{X}_1} \Big(\int_{\mathfrak{X}_2} f(x_1, x_2) \, \mathrm{d}\mu_2(x_2) \Big) \, \mathrm{d}\mu_1(x_1)$$

$$= \int_{\mathfrak{X}_2} \Big(\int_{\mathfrak{X}_1} f(x_1, x_2) \, \mathrm{d}\mu_1(x_1) \Big) \, \mathrm{d}\mu_2(x_2)$$

Useful results from measure theory

Exchange of limit and integration: $(f_n)_{n\in\mathbb{N}}$ m$\underline{\text{able}}$ f$\underline{\text{tions}}$ on $\mathfrak{X}$

- Monotone convergence theorem

  f_n non-negative, and $f_n \uparrow f$

- Dominated convergence theorem

  $|f_n| \le g$ for some $g \in L^1(\mu)$, and $f_n \to f$

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n \big)\, \mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n \, \mathrm{d}\mu \Big)$$

Exchange of order of integrations: $f \colon \mathfrak{X}_1 \times \mathfrak{X}_2 \to \mathbb{R}$ m$\underline{\text{able}}$ f$\underline{\text{tion}}$

μ_1 and μ_2 two (σ-finite) measures on $\mathfrak{X}_1$ and $\mathfrak{X}_2$, respectively

- Fubini's theorem

  $f \ge 0$ or $f \in L^1(\mu_1 \otimes \mu_2)$

$$\Rightarrow \quad \int_{\mathfrak{X}_1 \times \mathfrak{X}_2} f \, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{\mathfrak{X}_1} \Big(\int_{\mathfrak{X}_2} f(x_1, x_2)\, \mathrm{d}\mu_2(x_2) \Big) \mathrm{d}\mu_1(x_1)$$

$$= \int_{\mathfrak{X}_2} \Big(\int_{\mathfrak{X}_1} f(x_1, x_2)\, \mathrm{d}\mu_1(x_1) \Big) \mathrm{d}\mu_2(x_2)$$

Useful results from measure theory

- Fatou's lemma

  $(f_n)_{n\in\mathbb{N}}$ non-neg. m$\underline{\text{able}}$ f$\underline{\text{tions}}$ $\Rightarrow$ $\int(\liminf f_n) \leq \liminf(\int f_n)$

Exchange of limit and integration: $(f_n)_{n\in\mathbb{N}}$ m$\underline{\text{able}}$ f$\underline{\text{tions}}$ on $\mathfrak{X}$

- Monotone convergence theorem

  f_n non-negative, and $f_n \uparrow f$

- Dominated convergence theorem

  $|f_n| \leq g$ for some $g \in L^1(\mu)$, and $f_n \to f$

$$\Rightarrow \quad \int_{\mathfrak{X}} \big(\lim_{n\to\infty} f_n\big)\, \mathrm{d}\mu = \lim_{n\to\infty} \Big(\int_{\mathfrak{X}} f_n \, \mathrm{d}\mu \Big)$$

Exchange of order of integrations: $f: \mathfrak{X}_1 \times \mathfrak{X}_2 \to \mathbb{R}$ m$\underline{\text{able}}$ f$\underline{\text{tion}}$

μ_1 and μ_2 two (σ-finite) measures on $\mathfrak{X}_1$ and $\mathfrak{X}_2$, respectively

- Fubini's theorem

  $f \geq 0$ or $f \in L^1(\mu_1 \otimes \mu_2)$

$$\Rightarrow \quad \int_{\mathfrak{X}_1\times\mathfrak{X}_2} f \, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{\mathfrak{X}_1} \Big(\int_{\mathfrak{X}_2} f(x_1, x_2)\, \mathrm{d}\mu_2(x_2) \Big) \mathrm{d}\mu_1(x_1)$$
$$= \int_{\mathfrak{X}_2} \Big(\int_{\mathfrak{X}_1} f(x_1, x_2)\, \mathrm{d}\mu_1(x_1) \Big) \mathrm{d}\mu_2(x_2)$$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\qquad X\colon \Omega \to \mathbb{R}$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\qquad X : \Omega \to \mathbb{R}$

Distribution of X: $\mu = \mu_X$

the probability measure on $\mathbb{R}$ given by

$$\mu[B] = \mathrm{P}[X \in B] \qquad\qquad (B \subset \mathbb{R} \text{ Borel set})$$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\qquad X\colon\Omega\to\mathbb{R}$

Distribution of X: $\mu = \mu_X$

the probability measure on $\mathbb{R}$ given by

$$\mu[B] = \mathrm{P}[X \in B] \qquad\qquad (B \subset \mathbb{R} \text{ Borel set})$$

Cumulative distribution function (cdf): $F = F_X$

$F\colon\mathbb{R}\to[0,1]$ given by $F(x) = \mathrm{P}[X \le x]$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\quad X : \Omega \to \mathbb{R}$

Distribution of X: $\mu = \mu_X$

the probability measure on $\mathbb{R}$ given by

$$\mu[B] = \mathrm{P}[X \in B] \qquad (B \subset \mathbb{R} \text{ Borel set})$$

Cumulative distribution function (cdf): $F = F_X$

$F : \mathbb{R} \to [0, 1]$ given by $F(x) = \mathrm{P}[X \leq x]$

- F is non-decreasing
- $F(x) \searrow 0$ as $x \searrow -\infty$, and $F(x) \nearrow 1$ as $x \nearrow +\infty$
- F is right continuous: $F(x) \searrow F(x_0)$ as $x \searrow x_0$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\quad X: \Omega \to \mathbb{R}$

Distribution of X: $\mu = \mu_X$

the probability measure on $\mathbb{R}$ given by

$$\mu[B] = \mathsf{P}[X \in B] \qquad (B \subset \mathbb{R} \text{ Borel set})$$

Cumulative distribution function (cdf): $F = F_X$

$F: \mathbb{R} \to [0,1]$ given by $F(x) = \mathsf{P}[X \leq x]$

- F is non-decreasing
- $F(x) \searrow 0$ as $x \searrow -\infty$, and $F(x) \nearrow 1$ as $x \nearrow +\infty$
- F is right continuous: $F(x) \searrow F(x_0)$ as $x \searrow x_0$

Characteristic function: $\chi = \chi_X$

$\chi: \mathbb{R} \to \mathbb{C}$ given by $\chi(\theta) = \mathsf{E}[e^{i\theta X}]$

Distribution of real-valued random variable

$\mathbb{R}$-valued random variable $\qquad X\colon \Omega \to \mathbb{R}$

Distribution of X: $\mu = \mu_X$

the probability measure on $\mathbb{R}$ given by

$$\mu[B] = \mathsf{P}[X \in B] \qquad (B \subset \mathbb{R} \text{ Borel set})$$

Cumulative distribution function (cdf): $F = F_X$

$F\colon \mathbb{R} \to [0,1]$ given by $F(x) = \mathsf{P}[X \le x]$

- F is non-decreasing
- $F(x) \searrow 0$ as $x \searrow -\infty$, and $F(x) \nearrow 1$ as $x \nearrow +\infty$
- F is right continuous: $F(x) \searrow F(x_0)$ as $x \searrow x_0$

Characteristic function: $\chi = \chi_X$

$\chi\colon \mathbb{R} \to \mathbb{C}$ given by $\chi(\theta) = \mathsf{E}[e^{\mathrm{i}\theta X}]$

- χ is continuous and $\chi(0) = 1$
- $|\chi(\theta)| \le 1$ and $\chi(-\theta) = \overline{\chi(\theta)}$

Independence

Independence: abbreviated $\perp\!\!\!\perp$

Independence

Independence: abbreviated $\perp\!\!\!\perp$

sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \ldots \subset \mathcal{F}$

- $\perp\!\!\!\perp$ iff for all $G_1 \in \mathcal{G}_1, G_2 \in \mathcal{G}_2, \ldots, G_n \in \mathcal{G}_n$ we have $P[G_1 \cap G_2 \cap \cdots \cap G_n] = \prod_{j=1}^{n} P[G_j]$

Independence

Independence: abbreviated $\perp\!\!\!\perp$

sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \ldots \subset \mathcal{F}$

- $\perp\!\!\!\perp$ iff for all $G_1 \in \mathcal{G}_1$, $G_2 \in \mathcal{G}_2, \ldots, G_n \in \mathcal{G}_n$ we have $\mathrm{P}[G_1 \cap G_2 \cap \cdots \cap G_n] = \prod_{j=1}^{n} \mathrm{P}[G_j]$

random variables $X_1, X_2, X_3, \ldots$

- $\perp\!\!\!\perp$ iff the sigma-algebra generated by them are $\perp\!\!\!\perp$

Independence

Independence: abbreviated $\perp\!\!\!\perp$

sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \ldots \subset \mathcal{F}$

- $\perp\!\!\!\perp$ iff for all $G_1 \in \mathcal{G}_1$, $G_2 \in \mathcal{G}_2, \ldots, G_n \in \mathcal{G}_n$ we have $P[G_1 \cap G_2 \cap \cdots \cap G_n] = \prod_{j=1}^n P[G_j]$

random variables $X_1, X_2, X_3, \ldots$

- $\perp\!\!\!\perp$ iff the sigma-algebra generated by them are $\perp\!\!\!\perp$

events $E_1, E_2, E_3, \ldots \in \mathcal{F}$

- $\perp\!\!\!\perp$ iff their indicators are $\perp\!\!\!\perp$