randomness

Stengel §2.4

goal: mathematical representation of "random" variables, vectors, signals & their statistics

* the universe may – or may not – be "random"; this epistimological issue doesn't change the fact that "randomness" is a useful fiction

– you may regard "randomness" as arising from unmodeled phenomena or uncertainty (i.e. lack of complete knowledge) in models

def: a random variable is a function $X : \Omega \to \mathbb{R}$ over a sample space Ω

*note that there is nothing "random" about X; all the "uncertainty" lies in probability $P(W)$ of event $W \subset \Omega$

if $|\Omega| < \infty$ easy to define $P(W)$:

- if $|\Omega| < \infty$, easy to define $P(W)$: each $\omega \in \Omega$ gets probability "mass" $P(\omega)$, and $P(W) = \sum_{w \in W} P(w)$

→ what is domain & range of P?

– $P : 2^{\Omega} \to [0,1]$

(given sets A & B, overload exponent notation $B^A = \{ f : A \to B \}$)

– P termed a <u>probability mass function</u> (pmf)

- can compute prob of x giving outcome $\xi \in \mathbb{R}$:

$$P_x(\xi) = \sum \{ P(\omega) \mid \omega \in \Omega, x(\omega) = \xi \}$$

→ does $P = P_x$?

(what kind of object is P_x?)

– no: $P_x : \mathbb{R} \to [0,1]$

<u>ex:</u> suppose that, after shuffling a

ex: suppose that, after shuffling a standard deck of playing cards, each sequence has equal probability

→ when you draw top card, what is:

- sample space Ω? $|\Omega| = 52$
- $P(A)$? $4/52$ - $P(\text{spades})$? $1/4$ - $P(A \text{ of spades})$? $1/52$

→ when you draw a "hand" (i.e. 5 cards):

- sample space Ω? $|\Omega| = \binom{52}{5}$
- $P(AAAA)$? - $P(\text{flush})$? (all same suit)

→ if I reshuffle the deck, do answers change?

- no; we're given that all seq. equally likely

∘ if $|\Omega| = \infty$, e.g. $\Omega = \mathbb{R}^n$, pathologies can arise (cf Banach-Tarski paradox)

~ easier if $\Omega = \mathbb{R}^n$ and $\exists\, p : \Omega \to \mathbb{R}$ s.t.

$$P(W) = \int_W p(\xi)\, d\xi$$

p a probability density function (pdf)

- main example is Gaussian (or normal) rv

$$\forall W \subset \mathbb{R}^n : P(W) = \int_W \frac{\exp\left(-\frac{\sigma}{2}(\xi - \mu)^2\right)}{\sqrt{2\pi\sigma^2}}\, d\xi$$

- given rv $x : \Omega \to \mathbb{R}$, several statistics we may wish to compute:
  - a statistic is a function that inputs rv's and outputs numbers

def: expectation or mean:

$$E[x] = \sum_{\omega \in \Omega} x(\omega) \cdot P(\omega), \quad |\Omega| < \infty$$

- if $\Omega = \mathbb{R}^n$ & x has density $p : \mathbb{R}^n \to \mathbb{R}$:

$$E[x] = \int x \cdot p(x)\, dx, \quad P(W) = \int_W p(\xi)\, d\xi$$

def: variance:

$$\mathrm{Var}[x] = \sum_{\omega \in \Omega} \left(x(\omega) - E[x]\right)^2 \cdot P(\omega)$$

$$\text{or} = \int (\xi - E[x])^2 p(\xi)\, d\xi$$

ex: Gaussian (or normal) has

mean μ, variance σ^2

- if r.v.'s $x, y : \Omega \to \mathbb{R}$ defined over same sample space, can compute joint probabilities:

$$P_{xy}(\xi, \eta) = \sum \{ P(\omega) : x(\omega) = \xi, y(\omega) = \eta \}$$

→ what kind of object is P_{xy}?

- $P_{xy} : \mathbb{R} \times \mathbb{R} \to [0,1]$

- if x & y are independent,

then $P_{xy}(\xi, \eta) = P_x(\xi) P_y(\eta)$

def: conditional probability

$$P_{x|y}(\xi | \eta) = \frac{P_{xy}(\xi, \eta)}{P_y(\eta)}$$

- Bayes' rule follows:

$$P_{x|y}(\xi | \eta) = \frac{1}{P_y(\eta)} P_{y|x}(\eta | \xi) P_x(\xi)$$

∘ given finite collection of r.v.'s $\{X_k:\Omega\to\mathbb{R}\}_{k=1}^{n}$,

can regard as random vector $X:\Omega\to\mathbb{R}^n$

- joint probability $P_X(\xi)=P_{X_1\cdots X_n}(\xi_1,\cdots,\xi_n)$
- expectation $E[X]\in\mathbb{R}^n$, $[E[X]]_k=E[X_k]$
- covariance $E[(X-E[X])(X-E[X])^T]$

→ what kind of object is covariance?

- $E[(X-E[X])(X-E[X])^T]\in\mathbb{R}^{n\times n}$

ex: Gaussian (or normal) rv

w/ mean $\mu\in\mathbb{R}^n$, covariance $\Sigma\in\mathbb{R}^{n\times n}$:

$$\forall W\subset\mathbb{R}^n: P(W)=\int_W \frac{\exp\left(-\frac{1}{2}(\xi-\mu)^T\Sigma(\xi-\mu)\right)}{\sqrt{(2\pi)^n|\Sigma|}}\,d\xi$$

* if $X:\mathbb{R}^n\to\mathbb{R}^n$ is a Gaussian random vector,

$E[X]=\mu\in\mathbb{R}^n$, $Cov[X]=\Sigma\in\mathbb{R}^{n\times n}$,

and $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$ are given, then $y = Ax + b$ is a Gaussian rv,

$$E[y] = A\mu + b, \quad Cov[y] = A \Sigma A^T$$

→ this is a very special property not satisfied by most random variables!

def: random process $X : \Omega \times T \to \mathbb{R}$ is a collection of rvs $\{X_t : \Omega \to \mathbb{R}\}$ indexed by time $t \in T$ ($T = \mathbb{Z}$ or $\mathbb{R}$)