L02. Probability Review

$\Omega = \{\xi_1, \xi_2, \ldots, \xi_n\}$

Ω: set of certain events (outcomes from an experiment) or probability space

a : specific event, any subset of Ω

$a \subseteq \Omega$, $a \in \mathcal{F}$ (Field: all subsets of Ω)

Ex: Toss a coin

$\Omega = \{h, t\}$

$a_1 = h$, $a_2 = t$, $a_3 = \Omega$

Q: is $a_4 = 0$ ev

Ex: $\Omega = \{$'roll 1', 'roll 2', 3, 4, 5, 6$\}$ on a dice.

$a_1 =$ '1' , $a_2 =$ '1' $\cup$ '2' , etc.

Ex2: Measuring time a laser beam 'flights' to an obstacle and comes bac

$\Omega = \{t \mid t \in (0, T)\}$

$a_1 = T/2$, $a_2 = \{t \mid t \in (t_1, t_2), 0 < t_1 < t_2 < T\}$

☆ A. Axiomatic definition of probability. (Axiom = A proposition that commends itself to general acceptance. A simple proposition).

I) $p(a) \geq 0$

II) $p(\Omega) = 1$

(set operations) $\boxed{\Omega \; a \; b}$

III) if $a \cap b = 0$ (mutually exclusive events) then $p(a \cup b) = p(a) + p(b)$

(Kolmogoroff's work)

B. Relative-frequency def.

$p(a) = \lim_{n \to \infty} \frac{n_a}{n}$

Deduced by observation. I, II compatible

III problematic

C. Classical def.

$p(a) = \frac{N_a}{N}$

Deduced logically.

Some events 'a' and prob. spaces Ω are hard to count.

Ex: 2 dices sum 7. $1/11$? (sums) $3/21$? (pairs) $\boxed{6/36}$? (pairs exhaustive)

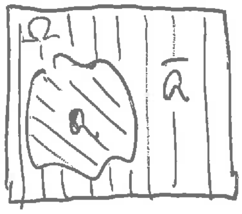
D. Probability as a Measure of Belief

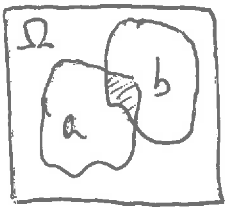
Inductive reasoning: "it is possible that X is guilty or not"

Subjective Knowledge is better than nothing.

Bayes Theory.

Corollaries from Axioms





$a \cap b$

$p(0) = 0$

$p(a) = 1 - p(\bar{a}) \leq 1$

if $a \cap b \neq 0$ then $p(a \cup b) = p(a) + p(b) - p(a \cap b)$

Proof: Since the null element 0 yields $a \uplus 0 = a$, $a \cap 0 = 0$

$$p(a) = p(a \uplus 0) = p(a) + p(0) \Leftrightarrow p(0) = 0$$

Similarly $a + \bar{a} = \Omega$ $\quad a \cap \bar{a} = 0$

$$\Rightarrow p(a) + p(\bar{a}) = p(\Omega) \overset{(A.II)}{=} 1$$

and $a \cup b = a \cup (\bar{a} \cap b)$, $\quad b = (a \cap b) \cup (\bar{a} \cap b)$

$$\left.\begin{array}{l} \Rightarrow p(a \cup b) = p(a) + \underbrace{p(\bar{a} \cap b)} \\ p(b) = p(a \cap b) + p(\bar{a} \cap b) \end{array}\right\} = p(a) + p(b) - p(a \cap b)$$

Def: Experiment. : $\Omega = \{\xi_1, \xi_2, \dots, \xi_n\}$ outcomes of the experiment.

$F = \{$ all certain subsets of $\Omega\}$ (Borel field)

$p(a)$ $\quad a \in F$. $\quad$ event a belongs to the field F.

then an experiment is $E: (\Omega, F, p)$

Ex: experiment of rolling a dice.

$\Omega = \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6\}$ $\quad \xi_i$ outcome of rolling 'i' on the dice.

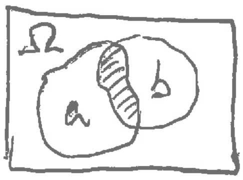
$$F = \{0, \underbrace{\xi_1, \xi_2, \dots}_{\text{single}}, \underbrace{\xi_1 \cup \xi_2, \xi_1 \cup \xi_3, \dots}_{\text{pairs}}, \underbrace{\xi_1 \cup \xi_2 \cup \xi_3, \dots}_{\text{triples}}, \underbrace{\dots}_{\text{quad.}\dots}, \Omega\}$$

$p(\xi_i) = 1/6$

Q: define other experiments?

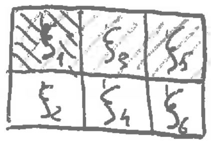
<u>Conditional</u> <u>probability</u>: Given two events a, b and $p(b)>0$

$$p(a|b) = \frac{p(a \cap b)}{p(b)} = \frac{p(a,b)}{p(b)}$$



Ex: rolling a dice, the probability of rolling '1' assuming an odd result.

Q: $p(\{\xi_1\} | \{odd\})$? $\{\text{'odd'}\} = \{\xi_1, \xi_3, \xi_5\}$



$$p(\xi_1 | \{odd\}) = \frac{p(\xi_1 \cap \{\xi_1, \xi_3, \xi_5\})}{p(\{\xi_1, \xi_3, \xi_5\})} \underset{A.III}{=} \frac{p(\xi_1)}{p(\xi_1)+p(\xi_3)+p(\xi_5)} = \frac{1/6}{3/6} = 1/3$$

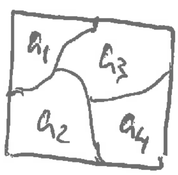
(Only using Axioms!)

Corollary: <u>Product Rule</u> $p(a,b) = p(a|b)\,p(b)$

<u>Total Probability</u> n mutually exclusive a_i \ $a_1 \cup a_2 \cup \cdots \cup a_n = \Omega$

$$p(b) = p(b|a_1)p(a_1) + \cdots + p(b|a_n)p(a_n)$$

$$= \sum_i^n p(b|a_i)p(a_i) \underset{\substack{\uparrow \\ Q?}}{=} \sum_i^n p(b, a_i)$$



(note: we'll see later for the cont. case and applied to pdf's)

Q: Ex: $b = a_4$

$$p(b|a_i) = 0 \qquad i = 1,2,3$$

$$p(b|b) = \frac{p(b)}{p(b)} = 1$$

$$p(b) = \sum_i^n p(b|a_i)p(a_i) = 0 + 0 + 0 + 1 \cdot p(a_4) = p(a_4)$$

Bayes Theorem

$$p(a_i \mid b) = \frac{\overbrace{p(b \mid a_i)\, p(a_i)}^{\text{product rule: } p(b, a_i)}}{\underbrace{p(b)}_{\text{total probability}}} = \frac{p(b \mid a_i)\, p(a_i)}{\sum_j^n p(b \mid a_j)\, p(a_j)} = \eta\, p(b \mid a_i)$$

Ex: a_i : 'I am ill' b : {'headache'} b' = {'headache' 'neck pain' 'fever'}

$p(b \mid a_i)$ easier to calculate. $\uparrow b \uparrow b'$

$p(a_i)$ prior $\bar{a}_i$ 'not ill' $a_i + \bar{a}_i = \Omega$

$p(b) = p(b \mid \text{'ill'})\, p(\text{'ill'}) + p(b \mid \text{'no ill'})\, p(\text{'no ill'})$

Independent events

$p(a, b) = p(a)\, p(b) \iff a, b$ indp. events.

Col: $p(a \mid b) = p(a)$ $p(b \mid a) = p(b)$

Ex: rolling a dice. $p(5_1, \Omega) = p(5_1)\, \underbrace{p(\Omega)}_{1}$

Ex: flipping 2 coins

$p(h, h) = p(h)\, p(h)$

$\Omega_2 = \{hh, ht, th, tt\}$

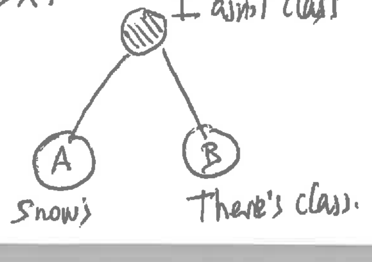
$\Omega_2 = \Omega_1 \times \Omega_1$

Conditional independent events:

$p(a, b \mid c) = p(a \mid c)\, p(b \mid c)$

Q: are a, b indep? $p(a, b) \neq p(a)\, p(b)$ not necessarily.

Ex:



C: today I haven't assisted class.

Q: Ω, F?? Introducing r.v.

Random variable (r.v.)

$X(\xi)$, $\xi \in \Omega$ experimental outcomes, but not nec. numbers

assigns a number to every outcome ξ

Ex: dice $\Omega = \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6\}$

r.v. $X(\xi_1) = 10$ $X(\xi_2) = 20$ $X(\xi_i) = 10i$

Ex: Tossing a coin. $\Omega = \{h, t\}$

r.v. X | $X(h) = 0$ $X(t) = 1$

Ex: Laser time-of-flight: $\Omega = \{t \mid t \in (0, T)\}$

r.v. $X(t) = \frac{c}{2} \cdot t$

note: subtle difference since Ω was already a subset of $\mathbb{R}$.
For most of this class, we need/work with infinite sets of events or 'continuous' events that easily translate to r.v.

probabilities → events. How to create events using r.v's ??

$X(\xi) \leq x$ conditions on X are events.

$\{X \leq x\}$

$\{x_1 \leq X \leq x_2\}$ Events defined by r.v's!!

Ex: dice.

ξ	$X(\xi)$
ξ_1	10
ξ_2	20
.	30
.	40
.	50
ξ_6	60

$\{X \leq 35\} = \{\xi_1, \xi_2, \xi_3\}$

Q: $\{20 \leq X \leq 35\} = \{\xi_2, \xi_3\}$

$\{X \leq 5\} = 0$

Distribution function

$p\{X \leq x\} = F_X(x)$ Non-decreasing function $F_X(x_1) \leq F_X(x_2)$, $\forall x$

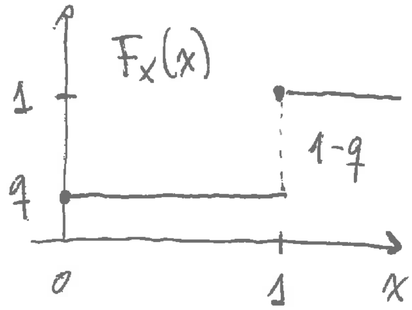
$\{X \leq +\infty\} = \Omega$ $F_X(+\infty) = p(\Omega) = 1$

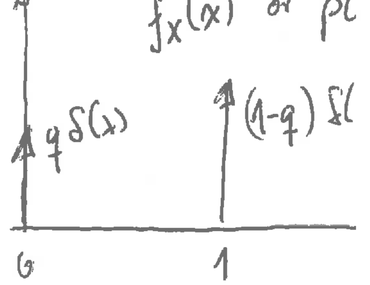
$\{X \leq -\infty\} = \emptyset$ $F_X(-\infty) = p(\emptyset) = 0$

Ex: Tossing a coin $\Omega = \{h, t\}$ $p(h) = q$ $p(t) = 1-q$

r.v. X $X(h) = 0$

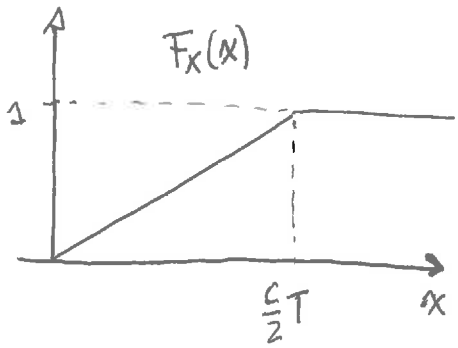
$X(t) = 1$

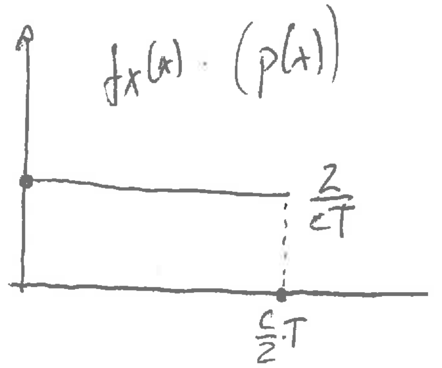




Ex: Laser time-of-flight $\Omega = \{t \mid t \in (0, T)\}$

r.v. $X(t) = \frac{c}{2}t$





Probability density function (pdf)

$$p(x) = f_X(x) = \frac{d F_X(x)}{dx}$$

Intuition: since $p\{x \leq X \leq x + \Delta x\} = \int_x^{x+\Delta x} f_X(x')\,dx'$ for Δx small

$$p\{x \leq X \leq x + \Delta x\} \simeq f_X(x) \cdot \Delta x$$

$$f_X(x) = \lim_{\Delta x \to 0} \frac{p\{x \leq X \leq x + \Delta x\}}{\Delta x}$$

$p(x) \geq 0$ (from $F_X(x)$ being non-decreasing) Q: $p(x) > 1$?

$\int_{-\infty}^{\infty} p(x)\,dx = F_X(\infty) - F_X(-\infty) = 1$

$F_X(x) = \int_{-\infty}^{x} p(x')\,dx'$

$P\{x_1 \leq X \leq x_2\} = \int_{x_1}^{x_2} p(x)\,dx$ (probability mass)

Q: $P\{X = x\} = 0$ is it true? Not always, for discrete distributions &
Always in the cont. case.

Multiple r.v. : Joint distribution function

$F_{XY}(x, y) = P\{X \leq x, Y \leq y\}$ Cartesian product of both Prob. sp

Joint ~~probability density function~~

$p(x, y) = f_{XY}(x, y) = \dfrac{\partial^2 F_{XY}(x, y)}{\partial x\,\partial y}$ \ $p(x, y) \geq 0$

$F_{XY}(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} p(x', y')\,dx'\,dy'$

Q: $F_{XY}(x, +\infty) \stackrel{?}{=} F_X(x)$
$F_{XY}(x, -\infty) \stackrel{?}{=} 0$

Note: The important thing here is that we already have discussed on probabilities for several events a, b. The introduction of r.v. was motivated to easily manipulate events on prob. space ($a \subseteq \Omega$) to numbers. and show how to create events from r.v., e.g. $\{X \leq x\}$

Now for multiples r.v. everything previously explained for events still holds, but we will skip the tedious math.

probabilities on events — distributions on r.v. — pdf on r.v.

$$p(a) \longrightarrow p\{X \leq x) = F_X(x) \xrightarrow{\frac{d}{dx}} \boxed{f_X(x) = p(x)}$$

Leap of faith here: redefinition of prev. concepts that hold for pdf's too

Product rule $\quad p(x,y) = p(x|y)\,p(y)$

Total probability $\quad p(x) = \int p(x,y)\,dy \quad$ (marginalization)

(conditioned) $\quad p(x|z) = \int p(x,y|z)\,dy \qquad$ Q: why? $x' = \{x|z\}$

Independent r.v.'s $\quad p(x,y) = p(x)\,p(y)$

Cond. Ind. r.v.'s $\quad p(x,y|z) = p(x|z)\,p(y|z)$

Bayes' Theorem $\quad p(x|y) = \dfrac{p(y|x)\,p(x)}{p(y)} = \dfrac{p(y|x)\,p(x)}{\int p(x,y)\,dx} = \eta\, p(y|x)$

(not depend on x)

Bayes conditioned $\quad p(x|y,z) = \dfrac{p(y|x,z)\,p(x|z)}{p(y|z)}$

Q: Prob Rob p. 14 is that a r.v.?