Probability - models of experiments with random outcomes.

Ingredients of a probability model

"Defn" of a probability space

- Sample space Ω
  Set of all possible outcomes
  Ex: roll a die
  $$\Omega = \{1, 2, 3, 4, 5, 6\}$$
  sample points.

- Events - the set of events - $\mathcal{F}$.
  A subset of Ω is called an event.
  $\mathcal{F}$ is a set of events.
  $\mathcal{F}$ is a set of subsets of Ω.

  $A = \{2, 4, 6\}$ - event that rolled an even number
  $$\mathcal{F} = \{A, \{1\}, \phi, \Omega, \dots\}$$

- Probability measure
  distribution
  P a function $: \mathcal{F} \longrightarrow \mathbb{R}$
  $A \in \mathcal{F}$, $P(A)$ - probability of A

P must satisfy
i) $0 \le P(A) \le 1 \quad \forall A \in \mathcal{F}$.
ii) $P(\Omega) = 1, \quad P(\phi) = 0$
iii) $\forall$ pairwise disjoint sets $A_1, A_2, \dots$
$$P(A_1 \cup A_2 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

Ω - set

$\mathcal{F}$ - set of subsets of Ω

P - fcn $: \mathcal{F} \to \mathbb{R}$

s.t. i) – iii)

$\underbrace{(\Omega, \mathcal{F}, P)}$

probability space.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \forall A, B \in \mathcal{F}.$$

Inclusion-exclusion principle

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i \cap A_j)$$

$$+ \sum_{1 \le i_1 < i_2 < i_3 \le n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots$$

$$+ (-1)^n P(A_1 \cap \dots \cap A_n).$$

Exercise: Show holds.

Ex: Roll of a die.

$\Omega = \{1,2,3,4,5,6\}$. $\mathcal{F}$ - set of all subsets of Ω

$P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}$

$P(\{1,2,3\}) = \frac{3}{6}$

$P(A) \quad \forall A \in \mathcal{F}$.

If Ω is finite, to specify P it is enough to specify the prob. of every outcome

$P(\omega) \quad \forall \omega \in \Omega$.

$\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$.

$P(\omega_1) + P(\omega_2) + \dots + P(\omega_k) = 1$.

Q: What if Ω is infinite?

Ex: Flip a coin until 1st heads.
Count the number of flips.
$\Omega = \{\infty, 1, 2, 3, 4, \ldots\}$
$\mathcal{F}$ - set of all subsets of Ω

$$P(\{k\}) = P(k) = P(\text{takes } k \text{ flips to get 1st H}) = \frac{1}{2^k}$$

$$1 = P(\Omega) = P(\infty) + \sum_{k=1}^{\infty} P(k) = P(\infty) + \underbrace{\sum_{k=1}^{\infty} \frac{1}{2^k}}_{1}$$

$P(\infty) = 0$.

$$P(\text{even number of tosses}) = \sum_{k=1}^{\infty} P(2k) = \sum_{k=1}^{\infty} \frac{1}{2^{2k}}$$

$$P(A) = \sum_{k \in A} P(k) \quad \forall A \in \mathcal{F}.$$

Random Variables

Def: A random variable is a function
$$X : \Omega \to \mathbb{R}$$

Ex: Roll 2 dice.
X_1 = value of 1st die
X_2 = " 2nd die
X_3 = Sum of the dice

Def: Let X be a RV. The prob. distr. of X is the collection of probabilities
$$P(X \in B) \text{ for "reasonable" sets } B \subseteq \mathbb{R}.$$

Ex: 1) Bernoulli distr.

A RV X has the Bernoulli distr w/ param p if

$$P(X=1)=p$$
$$P(X=0)=1-p.$$

2) Binomial distr params $n\in\mathbb{N}$, $p\in(0,1)$.

$$P(X=m)=\begin{cases}\binom{n}{m}p^m(1-p)^{n-m} & 0\le m\le n\\ 0 & \text{o/w}\end{cases}$$

Toss a Bernoulli(p) coin n times, count the number of heads $\sim \mathrm{Bin}(n,p)$.

3) Poisson distr. with param $\lambda>0$.

$$P(X=k)=\begin{cases}e^{-\lambda}\dfrac{\lambda^k}{k!}, & k\in\mathbb{Z}_{\ge 0}\\ 0 & \text{o/w}\end{cases}$$

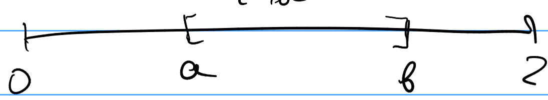
Check: $\displaystyle\sum_{k\ge 0}^{\infty} P(X=k) = 1.$

Ex: Pick a number uniformly at random from $[0,2]$.

$$P(X\in[a,b]) = \,?$$

$\underbrace{P(X=c)}_{=\alpha} = 0$ (Check)

$$P(X\in[a,b]) = \frac{b-a}{2} \neq \sum_{c\in[a,b]} P(X=c)$$



Defn: Let X be a RV. If a fcn f satisfies

$$P(X \le b) = \int_{-\infty}^{b} f(x)\,dx \quad \forall\, b \in \mathbb{R},$$

we say f is a density fcn (pdf) for X.

$$P(X \in B) = \int_B f(x)\,dx$$

Ex: $f(x) = \frac{1}{2} \mathbb{1}_{[0,2]}$ is a density fcn for Unif $[0,2]$.

$$P(X = c) = \int_c^c f(x)\,dx = 0.$$

Q: When is f a density fcn?
i) $f \ge 0$ ii) $P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f(x)\,dx = 1.$

Ex: Think of a RV which is neither discrete, nor has a density.

Defn: Let X be a RV. Its cdf $F_X(t)$ is

$$F_X(t) := P(X \le t) \quad \forall\, t \in \mathbb{R}$$

Check: If X has density $f_X(t)$ then

$$F_X'(t) = f_X(t).$$

Ex: A RV Z has the normal distr with parameters μ, σ^2, $Z \sim N(\mu, \sigma^2)$ if it has density

$$f_Z(t) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

The case when $\mu = 0$, $\sigma = 1$ called the standard normal

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Is f_Z a density?

$$\underbrace{f_Z \geq 0}, \quad \int_{\mathbb{R}} f_Z(t)\,dt = 1.$$

Change of var $- \mu = 0, \sigma^2 = 1.$

$$\underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1.}_{\text{Exercise.}}$$

Ex: A RV X has the exponential distr w/ param λ iff it has density

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & , t \geq 0 \\ 0 & , \text{o/w.} \end{cases}$$

Independence

Defn: Events A, B are indep if
$$P(A \cap B) = P(A)P(B)$$

Defn: Events $A_1, A_2, \dots, A_n$ indep if

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

$$\forall \quad i_1 < i_2 < \cdots < i_k$$

Defn: RVs $X_1, X_2, \dots$ are indep if
$\forall\ i_1 < i_2 < \dots < i_k \quad \forall\ B_1, B_2, \dots, B_k \subseteq \mathbb{R}$

the events $X_{i_1} \in B_1, \ X_{i_2} \in B_2, \dots, X_{i_k} \in B_k$
are independent.

Rmk: $A_1, \dots, A_n$ indep is not the same as saying A_i, A_j indep $\forall i \neq j$.

Exercise: Find such an example.

Ex: $X_1, X_2, \dots, X_n$ – indep Bernoulli(p)
$S_n = X_1 + X_2 + \dots + X_n \sim \text{Bin}(n, p)$

Check:

Expectation

Def: Let X be a discrete RV.
The expected value of X, EX is defined as

$$EX := \sum_t t\, P(X=t)$$

↑ possible values of X.

If X has density f, then

$$EX := \int_{-\infty}^{\infty} t\, f(t)\, dt$$

Thm: (The weak law of large numbers).
Let $X_1, X_2, X_3, \ldots$ be independent, identically distributed (i.i.d.) RVs.
Suppose $\mu = EX_i$ exists (& is finite)

$$S_n = X_1 + X_2 + \cdots + X_n$$

$\frac{S_n}{n}$ - average value

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

in probability.

$$\forall \varepsilon > 0 \qquad P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

Def: A seq of RVs $X_1, X_2, \ldots$ cv. in prob. to a RV X if $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0. \qquad X_n \xrightarrow{P} X$$

$X_n : \Omega \to \mathbb{R}$

$X : \Omega \to \mathbb{R}$

$\omega \in \Omega \qquad X_n(\omega) \longrightarrow X(\omega)$

<u>Defn:</u> A s'ce of RVs $X_1, X_2, \ldots$ cv to a RV X (defined on the same space) almost surely if

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

$$P(\{\omega \in \Omega : X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)\}) = 1.$$

$$X_n \xrightarrow{a.s.} X$$

<u>Defn:</u> A s'ce of RVs $X_1, X_2, \ldots$ cv to a RV X in distr if

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t) \quad \forall\, t \text{ where } F_X(t) \text{ is cts.}$$

$$X_n \xrightarrow{d} X$$

$$P(X_n \le t) \longrightarrow P(X \le t)$$

$\forall\, t \in \mathbb{R}$ where $F_X(t)$ is cts.

<u>Defn:</u> A s'ce of RVs X_n cv to X in L^p

$$\lim_{n\to\infty} E(|X_n - X|^p) = 0.$$

Thm: Suppose $X_1, X_2, \ldots, X$ are RVs defined on the same sp. Then

$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X$$

$$X_n \xrightarrow{L^p} X \implies$$

Exercise: For any implication not indicated, come up with examples of $X_1, X_2, \ldots X$ s.t. the implication is false.