Probability theory essentials

HA: 1.2.

Outline

- Events
- Probability and basic laws
- Random variables and distributions
- Moments and central moments
- Linear combination, minimum, maximum.

Next lecture

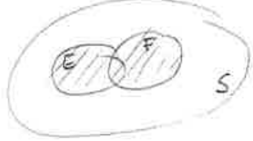
Z, Laplace, examples with Poisson, Exp, Erlang-k,

Sample spaces, Events,

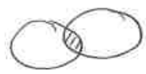
S : sample space : ~~poss~~ set of possible outcomes of an experiment
- discrete (e.g. die)
- continuous (e.g. life-time)

e : outcome of an experiment, $e \in S$
E : event, a subset of the sample space,

Combining events:



Union "or" $E \cup F = \{ e \in S \mid e \in E \vee e \in F \}$



Intersection: $E \cap F : \{ e \in S \mid e \in E \wedge e \in F \}$
"and"

Complement: $\bar{E} : \{ e \in S \mid e \notin E \}$

Probability

N : number of experiments
N(E) : number of outcomes in E

$$P(E) = \lim_{N \to \infty} \frac{N(E)}{N}$$

Axioms:

$0 \leq P(E) \leq 1$

$P(S) = 1$

if $\underbrace{E \cap F}_{\text{mutually exclusive}} = \emptyset$, then $P(E \cup F) = P(E) + P(F)$

Properties:

$P(\emptyset) = 0$

$P(\bar{E}) + P(E) = 1$

$P(E \cup F) = P(E) + P(F) - P(E \cap F)$

if $E \subseteq F$, then $P(E) \leq P(F)$

Conditional probability

Probability of event E, given that F has accured.

Def: $P(E|F) = \dfrac{P(E\cap F)}{P(F)}$ $\Rightarrow$ $P(E\cap F) = P(E|F)P(F)$

conditional probability

joint probability $(P(E,F))$

Example: E: { less than 4 with a die } = $\{1, 2, 3\}$

F : { even number with a die } = $\{2, 4, 6\}$

$$P(E|F) = \frac{1/6}{1/2} = 1/3$$

Independence

Events E and F are independent if and only if

Def: $P(E\cap F) = P(E)\cdot P(F)$

$$\Rightarrow P(E|F) = \frac{P(E\cap F)}{P(F)} = P(E)$$

Law of total probability

Complete set of mutually exclusive events: $\{E_1, E_2 \dots E_n\}$

partition of the sample space

- $\bigcup_i E_i = S$
- $E_i \cap E_j = \emptyset \quad \forall E_i, E_j \in S$

Then, for event A: $A = A\cap S = A\cap(\cup B_i) = \cup(A\cap B_i)$

$$P(A) = P(\bigcup_i (A\cap E_i)) = \sum_{i=1}^{n} P(A\cap E_i) = \sum_{i=1}^{n} P(A|E_i)P(E_i)$$

Bayes' formula

To calculate $P(E_i|F)$ when we know $P(F|E_j)$-s:

$$P(E_i|F) = \frac{P(E_i\cap F)}{P(F)} = \frac{P(F|E_i)P(E_i)}{\sum_{j=1}^{n} P(F|E_j)P(E_j)}$$

Bayes' example

Two dice: correct $P(1) = \frac{1}{6}$; faulty $P(1) = \frac{1}{2}$

Half of the dice on the market is faulty. You throw a 1. What is the probability that the die is faulty?

E_1: correct die $P(E_1) = \frac{1}{2}$

E_2 faulty die $P(E_2) = \frac{1}{2}$

F: through a 1 $P(F|E_1) = \frac{1}{6}$

$P(F|E_2) = \frac{1}{2}$

$$P(E_2|F) = \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{6}\cdot\frac{1}{2}} = \frac{\frac{1}{2}}{\frac{4}{6}} = \frac{1}{2}\cdot\frac{3}{2} = \boxed{\frac{3}{4}}$$

Random variables and distributions

Events do not always have an easy to capture notation

=>

Random variable X is a mapping: $X: \underbrace{S}_{\text{state space}} \to \underbrace{\mathbb{R}}_{\text{real numbers}}$

Often the mapping is trivial: eg., temperature, number of students...

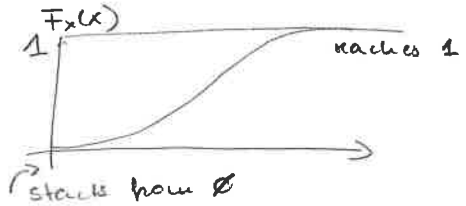
X < discrete < finite, infinite (discrete state space); continuous (continuous state space)

Cumulative distribution function (CDF) - to describe a random variable

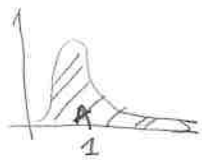
$F_X(x) = P(X \le x)$

↑ random variable: capital letter

$\bar{F}_X(x) = 1 - F_X(x) = P(X > x)$



Probability density function (PDF) - for continuous r.v.

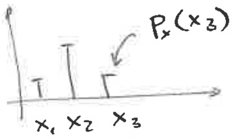


$$f_X(x) = \frac{dF_X(x)}{dx} = \lim_{dx \to 0} \frac{F_X(x+dx) - F_X(x)}{dx} = \lim \frac{P(x < X < x+dx)}{dx}$$

Probability mass function (PMF) - for discrete r.v.

$p_X(x) = P(X = x)$

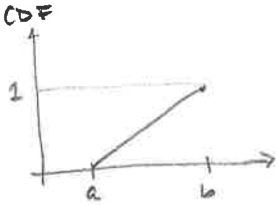
$\sum_i p_X(x_i) = 1$



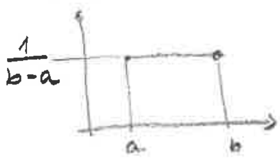
Simple example

- Continuous, uniform distribution: $U[a, b]$

CDF



PDF



Moments and central moments

1st moment: Expectation: $\bar{x} = E[X] = \begin{cases} \int_{-\infty}^{\infty} x f(x)\,dx \\ \sum_i x_i p_i \end{cases}$

nth moment: $E[X^n] = \begin{cases} \int_{-\infty}^{\infty} x^n f(x)\,dx \\ \sum_i x_i^n p_i \end{cases}$

Variance (second central moment) $\sigma^2 = Var[X] = E[(X-\bar{X})^2] = \int_{-\infty}^{\infty} (x-\bar{x})^2 f(x)\,dx = E[X^2] - E[X]^2$

↳ standard deviation $\sigma = \sqrt{Var[X]}$

Skewness (third central moment) ...

Eg. $U[a,b]$: $E[X] = \int_a^b x \frac{1}{b-a}\,dx = \frac{1}{b-a} \cdot \frac{1}{2}\left[x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$, etc.

Some useful properties

$E[cX] = c\,E[X]$ why? $E[cX] = \int_{-\infty}^{\infty} c x f(x)\,dx = c\int_{-\infty}^{\infty} x f(x)\,dx \dots$

$E[X_1 + X_2 \dots] = E[X_1] + E[X_2] \dots \rightarrow$

Note! $E[(X_1+X_2)^2] \neq E[X_1^2] + E[X_2^2]$!

$Var[cX] = c^2\,Var[X]$

$Var[X_1 + X_2 \dots] = Var[X_1] + Var[X_2]$ only if X_1 and $X_2 \dots$ are independent.

E.g.: What is the expected value of trowing the dice trice?

$E[X_1] = 3.5$, $E[X_2] = 3.5$, $E[X_1+X_2] = 7$

Combination of random variables

$f(x) = \alpha_1 f_1(x_1) + \alpha_2 f_2(x)$ $\quad \alpha_1 + \alpha_2 = 1$

$E[X] = \alpha_1 E_1[X] + \alpha_2 E_2[X]$

$E[X^2] = \alpha_1 E_1[X^2] + \alpha_2 E_2[X^2]$

Eg. 2 dice
- correct: $\{\frac{1}{6}, \frac{1}{6} \dots \frac{1}{6}\}$
- faulty: $\{\frac{1}{2}, 0, \dots \frac{1}{2}\}$

$\alpha_1 = \alpha_2 = 0.5$

$\Rightarrow f(x)$: $\{\frac{1}{3}, \underbrace{\frac{1}{12}, \dots \frac{1}{12}}_{4}, \frac{1}{3}\}$

Maximum of indep random variables

$P(\max(X_1 \dots X_n) \leq x) = P(X_1 \leq x, \dots X_n \leq x) = \prod P(X_i \leq x) = F_1(x) \cdot F_2(x) \dots$

Minimum

$P(\min(X_1 \dots X_n) > x) = \dots = \prod_i (1 - F_i(x))$

E.g. Through a dice 3 times. What is the probability that the maximum is less than 4 $\Rightarrow$ $P(\max < 4) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$!

If we have time

Famous distributions

Discrete: $\boxed{\text{Poisson distribution}}$

PMF:

$$P(X=k)=p_k=\frac{\lambda^k}{k!}e^{-\lambda} \qquad k=0,\dots,\infty$$

CDF: not nice

Mean: $\underline{\underline{E[X]}}=\sum_{k=1}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\cdot\lambda\underbrace{\sum_{k=0}^{\infty}\frac{\lambda^{k-1}}{k!}}_{e^{\lambda}}=\underline{\underline{\lambda}}$

$$\sum_{k=0}^{\infty}\frac{a^k}{k!}=e^a$$

Variance $Var[X]=\lambda$

Continuous: $\boxed{\text{Exponential}}$

PDF: $f(x)=\lambda e^{-\lambda x}, \quad x\geq 0$

CDF: $F(x)=1-e^{-\lambda x}$ ⇐ Derive at home

Mean: $E[X]=\int_0^{\infty} x\cdot e^{-\lambda x}\,dx=\dots=\frac{1}{\lambda}$ ⇐ Derive at home

$Var[X]=\frac{1}{\lambda^2}$

$\boxed{\text{Erlang-}k}$

Sum of k Exp. distributed rv: $X=X_1+X_2+X_k$, $X_i=Exp(\lambda)$

PDF: $f(x)=\frac{\lambda^k x^{k-1}}{(k-1)!}e^{-\lambda x}$ ⇐ Derive at home.

$E[X]=$ sum of id. r.v. $=\frac{k}{\lambda}$

$Var[X]=$ -"- $=\frac{k}{\lambda^2}$