Discrete Random Variables

- Usual : probability models in which a real number is assigned to each outcome in the sample space -- *the observation is called a* ***random variable X (r.v)***
- **DISCRETE R.V** : the range of X is a countable set $\boxed{S_X = \{x_1, x_2, x_3, \ldots\}}$
- Since S_X is countable, there is a 1-1 map onto the set of all positive integers.

Probability Mass Function (PMF)

- The PMF of a discrete r.v X
  $\boxed{P_X[x] = P[X = x]}$, as x varies
- i.e., $P_X(x)$ will be a series of δ-functions where there is a non-zero probability that X=x. The height of these δ-functions gives the probability that X=x

PMF (cont.)

- Th. 2.1 : For discrete r.v X, with PMF $P_X(x)$ and range $S_X=\{x_1, x_2, \ldots\}$

  (a) $\boxed{P_X[x] \geq 0}$: with $\boxed{P_X[x] = p_i\ \delta(x - x_i)}$

  (b) $\boxed{\sum_i p_i = 1}$

  (c) For any subs $\boxed{B \subseteq S_X}$ $\boxed{P[B] = \sum_{j \in B} p_j}$

Useful Discrete R.V's

- **BERNOULLI** : PMF

$$P_X[x] = (1-p)\,\delta(x) + p\,\delta(x-1)$$

- - experiments with only 2 possible outcomes
- p - parameter (0<p<1)

Useful Discrete R.V's

- **GEOMETRIC** : PMF

$$P_X(x) = p(1-p)^{x-1}\,.\,\delta(x-n)\ ,\ n = 1,2,\ldots$$

parameter p : 0<p<1

- check : *(normalization)*

$$\sum_n P_X(x) = p.\left[1 + (1-p) + (1-p)^2 + \ldots..\right]$$

infinite geometric series

$$= p.\left[\frac{1}{1-(1-p)}\right]$$ = 1

Useful Discrete R.V's

- **BINOMIAL :** PMF

$$P_X(x) = \binom{n}{x} p^x\,.(1-p)^{n-x}\ ,\quad x = 0,1,..,n$$

parameters 0 < p < 1 ; integer n$\geq$1

- n = # independent trials,
  p = success probability in a trial
  # successes in n trials = BINOMIAL r.v

Useful Discrete R.V's

- **PASCAL** : PMF

$$P_X(x) = \binom{x-1}{n-1} p^k . (1-p)^{x-k} \ , \quad x = k, k+1, \ldots$$

0<p<1 and integer k $\geq$ 1

- Pascal r.v = # trials needed to get the k^{th} success in n indep. trials

Useful Discrete R.V's

- **UNIFORM** : PMF
  X is uniformly distributed in integer interval between *k* and *m* :

$$P_X(x) = \begin{cases} \dfrac{1}{m-k+1} & , \ x = k, k+1, \ldots, m \\ 0 & , otherwise \end{cases}$$

  – N.B : the height of non-zero $P_X(x)$ is always the same ==> "uniform"

Useful Discrete R.V's

- **POISSON** : PMF
-
$$P_X(x) = \begin{cases} \alpha^x \dfrac{e^{-\alpha}}{x!} & , \ x = 0, 1, 2, \ldots.. \\ 0 & , \ otherwise \end{cases}$$
- α parameter
- suppose arrivals are indep. , the number of arrivals in a given time period is a Poisson r.v with probability of x arrivals in this time period.
  α = average number of arrivals in this time period.

CUMULATIVE DIST. FN. (CDF)

- CDF of r.v X is $F_X(x) = P[X \leq x]$
- $F_X(b) - F_X(a) = P[\ a < X \leq b]$, with b>a
- CDF for a discrete r.v has a staircase structure, with jumps of p_i at x_i where

$$P[X = x_i] = p_i$$

3 AVERAGES

- *mode* (x_{mod}) = number that occurs most frequently in the data set

  $P_X(x_{mod}) \geq P_X(x)$ for all x

- *Median* (x_{med}) = equal # of x's below x_{med} as above it : $P[X < x_{med}] = P[X > x_{med}]$

N.B : x_{mod} and x_{med} do NOT have to be unique for a given data set.

AVERAGES - the mean

- EXPECTED VALUE (mean) of r.v X

$$E[X] = \mu_X = \sum_{all\ x} x . P_X(x)$$

  - Bernoulli PMF : E[X] = p
  - Geometric PMF : E[X] = 1/p
  - Poisson PMF : E[X] = α
  - Binominal PMF : E[X] = n.p
  - Pascal PMF : E[X] = k/p
  - Uniform PMF : E[X] = (k+m)/2

Bernoulli Trials ==> Poisson PMF

- A large number (but *n* not known!) of Bernoulli trials (and also do NOT know *p,* the probability of success in a trial)

  Given the average number of successes in this large set of trials : α

  The Probability of k successes in these trials = $P(k) = \alpha^k e^{-\alpha}/k!$
  - a Poisson PMF

FUNCTIONS of a R.V

- Given a discrete r.v X, with PMF $P_X(x)$. For the derived r.v Y, with Y = g(X), its PMF is

$$\boxed{P_Y(y) = \sum_{x:g(x)=y} P_X(x)}$$

  – i.e., for given y, determine those x's s.t $g(x) = y$, then sum the corresponding $P_X(x)$

Expectation Values for Derived R.V

- Given r.v X with PMF $P_X(x)$, and Y=g(X)

  then $$\boxed{E[Y] = \mu_Y = \sum_{x \in S} g(x).P_X(x)}$$

- For any r.v X, $E[X - \mu_X] = 0$
  $E[a.X+b] = a.E[X] + b$

Variance & Standard Deviation

- Variance of a r.v X :

$$Var[X] = E\left[(X-\mu_X)^2\right] = E\left[X^2\right] - \{E[X]\}^2$$

- Standard Deviation : $\sigma_X = \sqrt{Var[X]}$
- Variances for common PMF's - Th 2.17

Conditional PMF

- For an event B which has P[B] > 0, the conditional PMF of X is

$$P_{X|B}(x) = P[X = x \mid B] = P_X(x) / P[B], \quad x \in B$$

- $$E[X \mid B] = \mu_{X|B} = \sum_{x \in B} x . P_{X|B}(x)$$
  $$E[g(X) \mid B] = \sum_{x \in B} g(x) . P_{X|B}(x)$$
  $$E[X] = \sum_i E[X \mid B_i] . P[B_i]$$