Probability Model

- as n gets large, the ratio $N_k(n)/n$ in the second expression approaches p_k
- the average number of active packets produced per 10 ms segment approaches

$$\langle A \rangle_n \rightarrow \sum_{k=0}^{48} k p_k \triangleq E[A]$$

- the expression on right hand side defined as expected value of A
- $E[A]$ is determined by the probabilities p_k
- it states the long-term average number of active packets produced per 10 ms period is $E[A]$
- the fraction of active packets that are discarded by the system in n trials is

$$\frac{\text{number of active packets discarded}}{\text{number of active packets produced}} = \frac{\sum_{k=M+1}^{48}(k-M)N_k(n)}{\sum_{k=0}^{48} kN_k(n)}$$

Probability Model

- $(k - M)$ is the number of packets that are discarded when $k > M$ active packets are produced
- divide numerator and denominator by n

$$\frac{\sum_{k=M+1}^{48}(k-M)N_k(n)/n}{\sum_{k=0}^{48} kN_k(n)/n} \rightarrow \frac{\sum_{k=M+1}^{48}(k-M)p_k}{\sum_{k=0}^{48} kp_k}$$

- right hand side is long-term fraction of active packets that are discarded

Probability Model

- target detection, target tracking, speech recognition, face recognition etc.
- signal received with unwanted signal (noise), signal to noise ratio SNR
- observed signal improves as the SNR increases, as the noise produces smaller perturbations about the desired signal
- resource sharing system
- computers, communication lines unsteady and random demand
- configure system such that demands of the users are met through the dynamic sharing of resources
- average response time as performance measure, queueing model for prediction of performance measures
- reliability of system, electronic transfer of fund, component failure rate

Statistical Learning

- goal of statistical learning theory is to study, in a statistical framework, the properties of learning algorithms
- most results take the form of so-called error bounds
- provide a framework for studying the problem of inference, that is of
- gaining knowledge, making predictions, making decisions or constructing models from a set of data
- there are assumptions of statistical nature about the underlying phenomena (in the way the data is generated)
- Nothing is more practical than a good theory - stated by Vapnik

Probability

- fair die - deduce probability of event even equals 3/6
- reasoning $P(1) = P(2) = \cdots = 1/6$ then $P(even) = 3/6$
- three postulates
  1. probability $P(A)$ of an event A is non-negative $P(A) \geq 0$
  2. certain event $P(S) = 1$
  3. if $A \cap B = \{\phi\}$ then $P(A \cup B) = P(A) + P(B)$
- relative frequency; $P(A)$ of an event A is the limit

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

- classical definition $P(A) = \frac{N_A}{N}$

Probability

- Bayes' rule
- let $B_1, B_2, \ldots, B_n$ be a partition of a sample space S
- the event A occurs; what is the probability of even B_i?

$$P[B_i|A] = \frac{P[A \cap B_i]}{P[A]} = \frac{P[A|B_i]P[B_i]}{\sum_{k=1}^{n} P[A|B_k]P[B_k]}$$

- in some random experiment in which the events of interest form a partition
- the "a priori probabilities" of these events $P[B_i]$ are the probabilities of the events before the experiment is performed

Probability

- Binary Communication system - inputs 0 or 1 into the system - transmitted receiver makes a decision about what was the input to the system based on the signal it received
- Say, user sends 0 with probability $1 - p$ and 1s with probability p and receiver makes random decision errors with probability ε

Probability

- B_1 be the event - receiver output was 1
- the probability of B_1 is

$$P[B_1] = P[B_1|A_0]P[A_0] + P[B_1|A_1]P[A_1] = \varepsilon\frac{1}{2} + (1-\varepsilon)\frac{1}{2} = \frac{1}{2}$$

- posteriori probabilities

$$P[A_0|B_1] = \frac{P[B_1|A_0]P[A_0]}{P[B_1]} = \frac{\varepsilon/2}{1/2} = \varepsilon$$

$$P[A_1|B_1] = \frac{P[B_1|A_1]P[A_1]}{P[B_1]} = \frac{(1-\varepsilon)/2}{1/2} = (1-\varepsilon)$$

- if ε is less than $1/2$, then input 1 is more likely than input 0 when a 1 is observed at the output of the channel
- independent event $P[A \cap B] = P[A]P[B]$