

Introduce myself, course rules.

Probability = Study of uncertain events that may be characterized mathematically.

Statistics = Science of determining probabilities from ~~data for actual outcome~~ ^{data}.

ex: We observe human behavior (gather statistics) and try to predict outcomes (behaviors) by assigning probabilities to future behaviors.

ex: Weather prediction (probabilities) based on past observations (statistics).

ex: Stock picking (prob of positive return) based on past quarterly results (statistics).

ex: Cell transmissions with noise from sun and other phones has (probability of) errors. Number of extra bits sent depends on (statistics of) noise.

ex: Quality control for integrated circuit production (probability of bit error in RAM chips) depends on (statistics of) particles in air in clean room. Mt. St. Helens eruptions wreaked havoc on Intel plant in Portland, Oregon.

ex: Digital processing of audio tapes (by selecting optimal filter) is a function of type of (i.e. statistics of) noise.

ex: Speech recognition is based on determining the most likely (highest probability) utterance heard given past experience (statistics of past speech).

ex. Maximum transmission rate for optical fiber (or any other communication channel) is given by Shannon bound: $C = B \log_2 \left(1 + \frac{P}{N}\right)$

where $C \equiv$ channel capacity in bits/sec

$B \equiv$ bandwidth of channel in Hz

$P \equiv$ power (ave) of transmitted signal in W

$N \equiv$ noise pwr (ave) in W

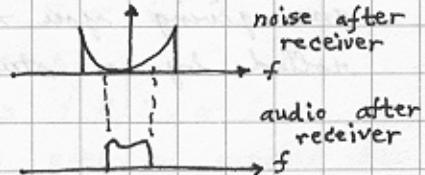
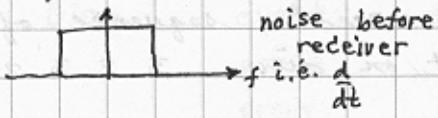
ex: Ultra-high bandwidth transmission increases B and reduces P . Other signals increase N . Because of \log_2 we are better off increasing B and reducing P .

ex: Decoding weak satellite transmissions may be accomplished by using Viterbi decoding to find the most likely (highest probability) sequence of bits sent given sparse codewords that are equally likely (i.e. have known statistics).

ex: Compression algorithms like jpeg use statistics of images to exploit unequal probabilities of images.

ex: Ethernet allows asynchronous communication on a single wire by delaying start times of packets by random amounts after a collision. Capacity \propto $\frac{1}{\text{prob of collision}}$ of network is determined by number of users and number of packets sent.

ex: FM has less noise than AM transmission by exploiting the uniformity of noise versus frequency. The FM signal requires filtering that changes the noise power spectral density ^{so} that the energy is pushed to frequencies that are filtered out.



ex: Semiconductor junctions are governed by statistical physics and the probability that electrons and holes are in particular energy states.

ex: Intel released a microprocessor with small arithmetic errors that would affect only one in 400,000,000 calculations and then only in the last few of many significant figures. They believed the probability of anyone finding the error was small enough that no one would notice it. It took only one week for a spreadsheet user to notice it... Why?

ex: Neuron firing patterns in the brain are somewhat random. The brain recognizes events by performing statistical pattern recognition. Information is encoded by intervals between 0.1V spikes called action potentials. Molecular noise affects spike timing.

ex: Random noise can improve searches for optimal solutions to problems. Simulated annealing is a search technique for finding the lowest point on a surface by searching at random and sometimes accepting worse solutions as you go.

ex: Mark Twain said: "There are lies, damned lies, and statistics." Before the course is over you will say "There are statistics, damned statistics, ... and lies." (on tests)

Illustration of probabilities: Can you tell whether I am giving you the correct sequence of numbers rolled by a student on die. 3 2 3 2 4 4 2 4