Assignment #3A

Probability Ideas, Confidence Intevals, Regression,

Graphing (Continued)

Sunday Sept 24, 11:55pm

Probability/Statistics: Questions

- FINISH UP FROM LAST TIME…
- If two RVs (A and B) are independent, what is P(A|B) in terms of P(A) and P(B)? What does the knowledge about the occurrence of B give you in this case?
- What information can you get from a CCDF that is not prominent in a pdf?
- What is the difference between mean, median and mode? When would you use each?
- How is CoV different from covariance and correlation coefficient?
- How are confidence intervals different from hypothesis tests?
- Why is the normal distribution so important?
- State one key implication of heavy-tailed distribution (in internet modeling). Why does poisson modeling fail for internet traffic?

Probability/Statistics: More Questions

- Under what conditions do binomial distributions tend towards normal distribution?
- How are the poisson and exponential distributions related?
- How are normal distributions “standardized”? What is the z-variate?
- Look up z < 0.45 in the normal distribution table:
  - http://www.math.unb.ca/~knight/utility/NormTble.htm
- In N(5,10), what is P(3.9 <= X <= 9.8) ?
- When do you use the t-distribution instead of the normal distribution for confidence intervals?
- What is the sampling distribution of the sample variance?
- What extra information does an interval estimate (like CI) give over a point estimate (mean) ?

Max-Likelihood Estimation Problem

- I have a biased coin that tosses Heads with (unknown probability) p
- I test the coin by tossing it 10 times.
- Data observed: 7 heads.
- What is the maximum likelihood estimate of p?
- Hint: see http://en.wikipedia.org/wiki/Maximum_likelihood

Confidence Intervals, Regression

- Given: n=6 random RTT samples (in ms):

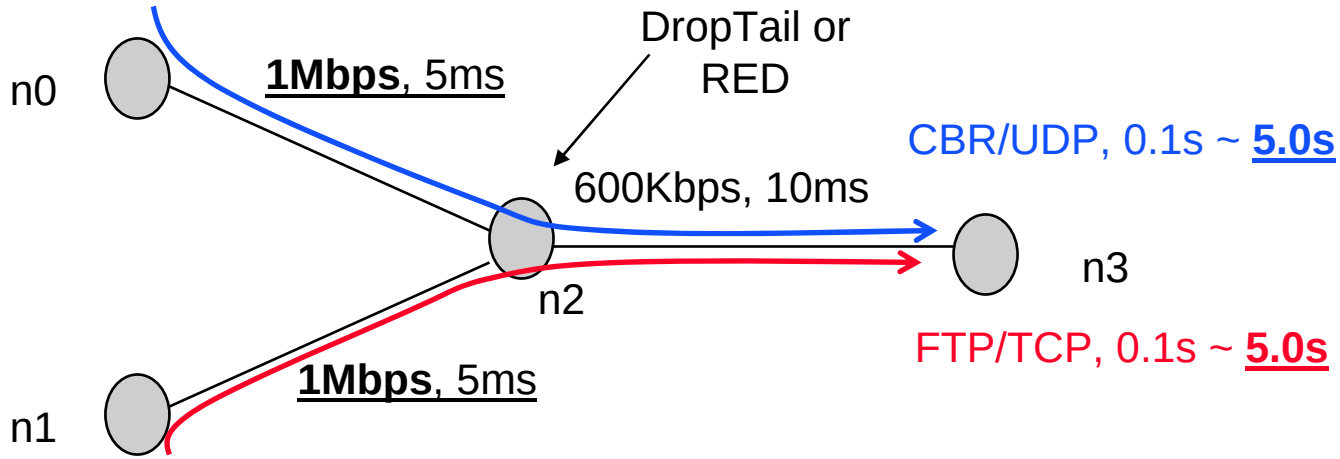
  {31, 23, 29, 52, 28, 41}

  Find: sample mean (xbar), sample standard deviation (s), 90% confidence interval (CI) & 99% CI for the *population* mean

- What is the coefficient of determination (R-squared): define? How is it different from the variance of the samples?
- What are SST, SSE, SSR? How are they related to each other and the variance of the samples?

Recall: Assignment #3: TCP

- TCP Dynamics



TCP Performance: Advanced Graphing

- Distribution of Performance
  - Graph the goodputs of all flows in a histogram.

- Based upon the lecture on graphing, come up with at least one other interesting view of TCP performance and graph it.

ALL Students

- Read the abstract/intro/conclusion of the two papers:
  - WiFi Rooftop Network Analysis paper and
  - BGP Instabilities paper
- Focus on the figures and see the “story they tell”…
- Write a brief summary about what are the interesting types of graphs used and why they are effective in making the points of the paper.

Graduate Students: Additional

- Read the abstract/intro/conclusion of the two papers:
  - VPN analysis paper
  - Faloutsos power laws paper
- Focus on the figures and see the “story they tell”…
- Write a brief summary about what are the interesting types of graphs used and why they are effective in making the points of the paper.

- Read the selected pages from “Internet Measurement” book [copies provided].
- Write a brief summary of ideas

Submission

- Write ns2 script to measure TCP (it is a TCP Tahoe) performance.

- Submissions:
  - Answers to probability questions
  - Ns2 simulation script;
  - All the required graphs and statistics.
  - All students: summary of graphing techniques in 2 papers (WiFi and BGP)
  - Grad students: summary of graphing techniques in 2 papers (VPN and Power Laws)

- Due **Sunday Sept 24, 11:55pm**

Note

- If you want to work on your own machine, you need to install ns-allinone-2.26 and graphing tool.
  - Talk to Neeraj (some of these versions may have changed)

- Example graph tool code (old version):
  - On your machine's directory
    ~/ns/ns-allinone-2.1b7/graph_v6.0.4/examples/
  - Downloadable at http://networks.ecse.rpi.edu/~harrisod
    which works with ns-2.1b5 (recommended) or ns-2.1b7-old