Probability with Engineering Applications
ECE 313 – Section C – Lecture 12

Lav R. Varshney

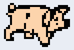
25 September 2017

Exams

- Oct 11, Wed, 8:45 - 10pm
  - Upto first five checkpoints
- Nov 15, Wed, 8:45 - 10pm
- Dec 18, Mon, 1:30 - 4:30pm

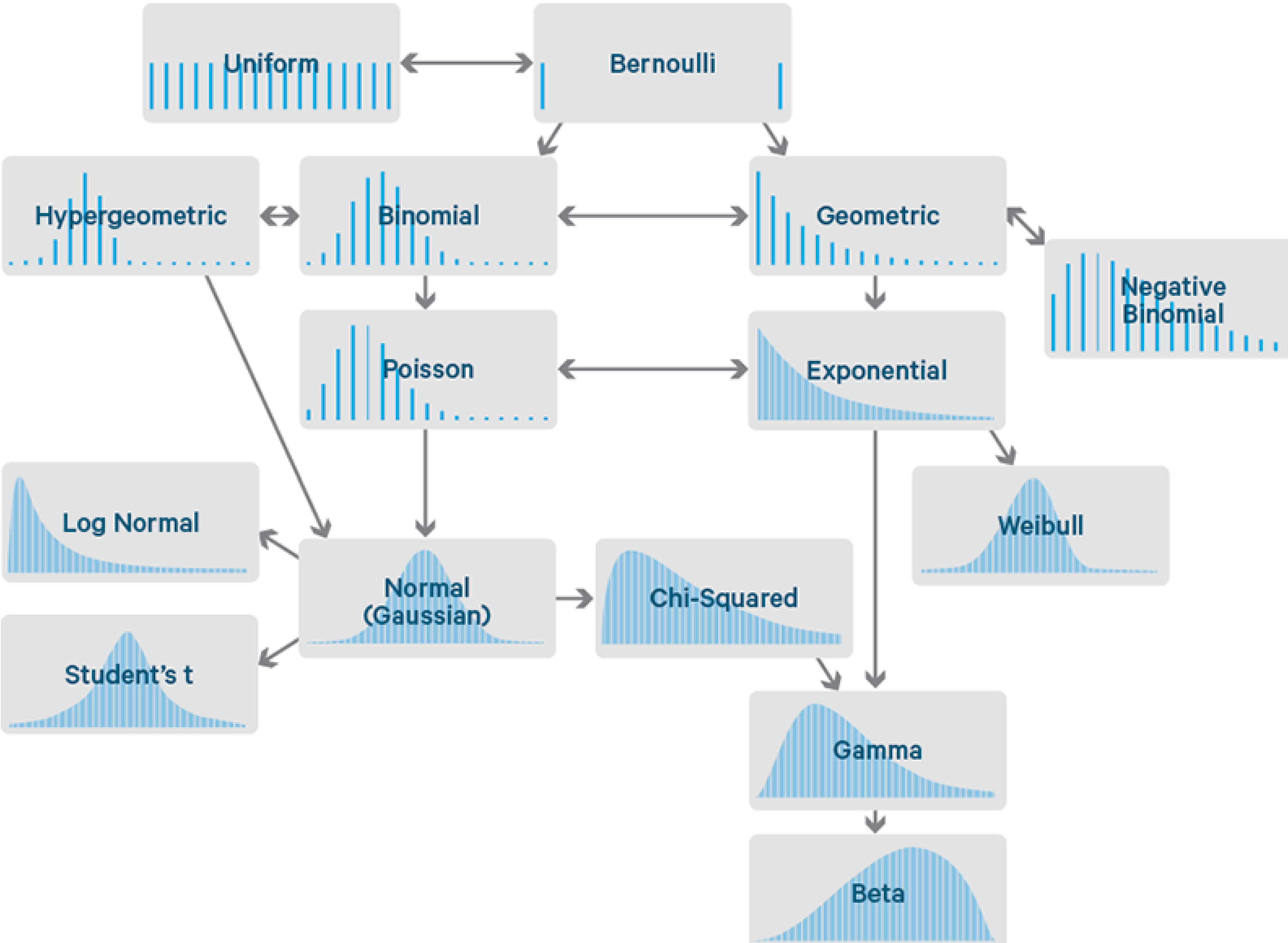
Non-parametric representation

The approximate relative frequencies of the various positions for a single pig, using a standardized surface, a trap-door rolling device, and a sample size of 11,954, are:

Position		Percentage
Side (no dot)		34.9%
Side (dot)		30.2%
Razorback		22.4%
Trotter		8.8%
Snouter		3.0%
Leaning Jowler		0.61%

[J. C. Kern, “Pig Data and Bayesian Inference on Multinomial Probabilities,” *Journal of Statistics Education*, vol. 14, no. 3, 2006.

Common probability distributions



[https://blog.cloudera.com/blog/2015/12/common-probability-distributions-the-data-scientists-crib-sheet/]

Parametric representation

Probability mass function	Parameter
Bernoulli	p
Binomial	n, p
Geometric	p
Poisson	λ

How do we find the values of the parameters from data?

Non-random parameter estimation

- Collect data and then estimate parameter using the observed data
- Suppose we decide an experiment is accurately modeled by a random variable X with pmf p_θ, where θ is an unknown (but not random) parameter
- When experiment is performed, suppose we observe a particular value k for X

Non-random parameter estimation

- According to probability model, the probability of k being the observed value for X (before experiment was performed), would have been $p_\theta(k)$
- It is said that the *likelihood* that $X = k$ is $p_\theta(k)$

Maximum likelihood estimation

- One standard approach for estimating non-random parameters is to maximize the likelihood of the observed data, by choosing the best value of the non-random parameter
- The MLE of θ for observation k is denoted $\hat{\theta}_{ML}(k)$ is the value of θ that maximizes the likelihood $p_\theta(k)$ with respect to θ

Maximizing expressions

- The best way we know to analytically maximize functions is to take the derivative with respect to the parameter we are maximizing, setting equal to zero, and solving.
- Often it may be easier to consider monotonic functions of the thing we care about, since the extremal value is the same.
  - Optimize log-likelihood, rather than likelihood directly, since log is a monotonic function

Example 1

- Suppose X has a Poisson distribution with some parameter $\lambda > 0$ which is unknown and a particular value k for X is observed.
- Find the maximum likelihood estimate, $\hat{\lambda}_{ML}$
- (Here λ plays the role of θ in the definition of ML estimation)

Example 2

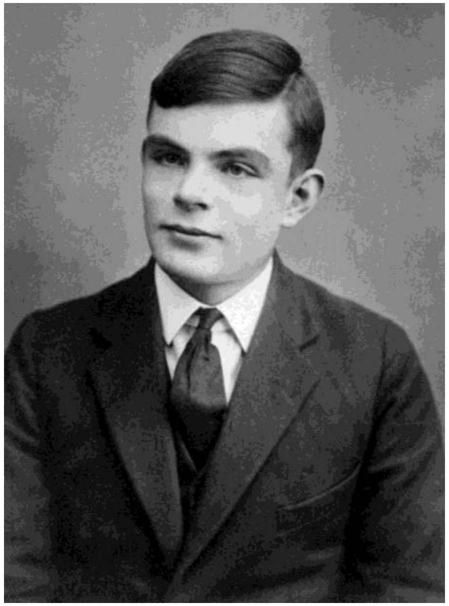
- Suppose X has the geometric distribution with some parameter p which is unknown and a particular value k for X is observed.
- Find the maximum likelihood estimate, $\hat{p}_{ML}$

Example 3

- Suppose we have a weighted coin, that shows heads with some unknown probability p each time it is flipped
- We flip it some large and known number n times and heads shows on k of the flips
- Find the ML estimate of p

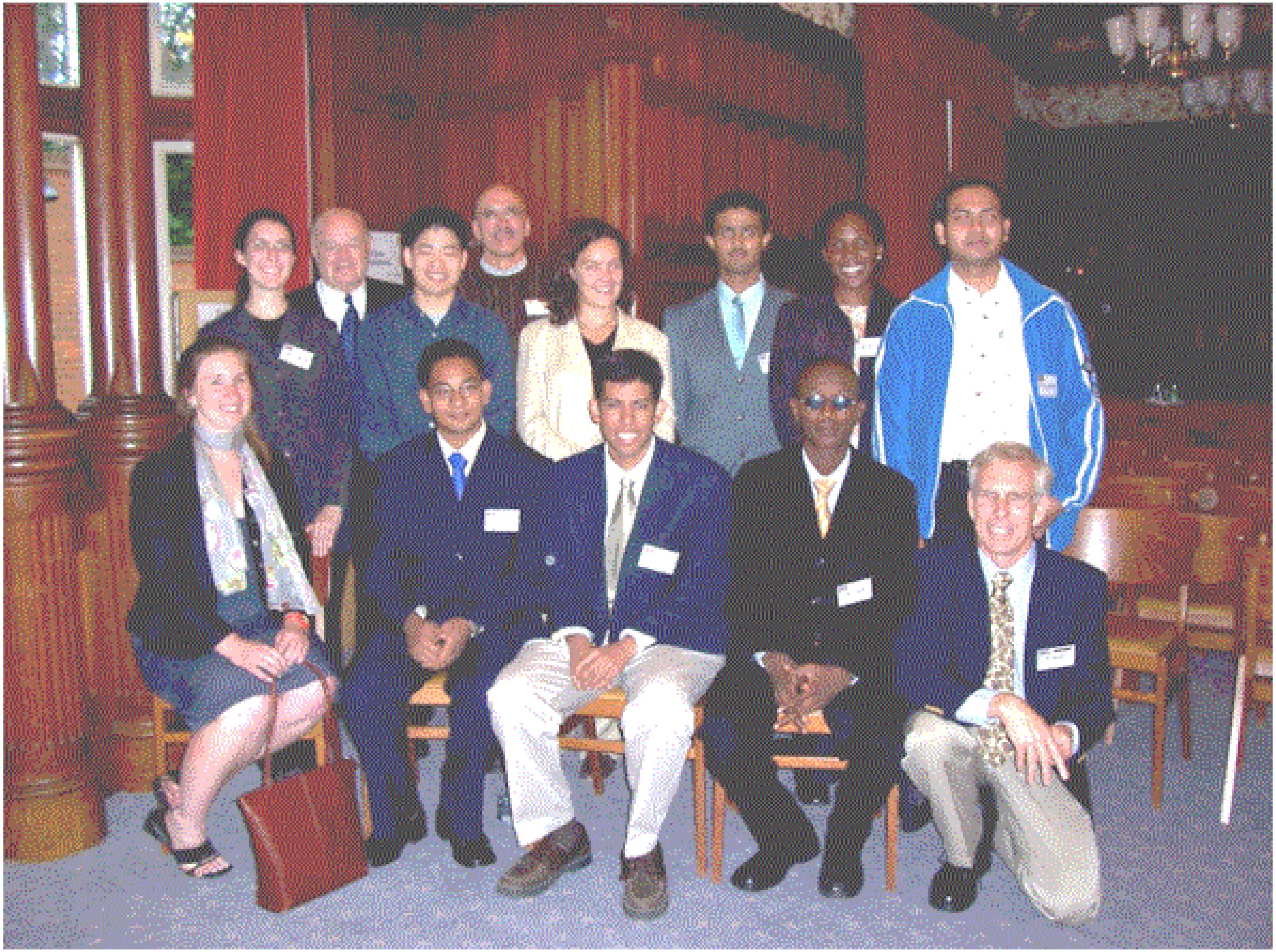
Unseen elements problem

Unseen elements problem



[https://vtspecialcollections.wordpress.com/2015/01/15/i-j-jack-good-virginia-techs-own-bletchley-park-connection/]

[http://aperiodical.com/2014/07/an-alan-turing-expert-watches-the-the-imitation-game-trailer/]



[http://ethw.org/2004_IEEE_Conference_on_the_History_of_Electronics]

Example 4

- Suppose X is drawn at random from the numbers 1 through n, with each possibility being equally likely but n is unknown
- We observe $X = k$
- Find $\hat{n}_{ML}$ as a function of k