Probability with Engineering Applications
ECE 313 – Section C – Lecture 11

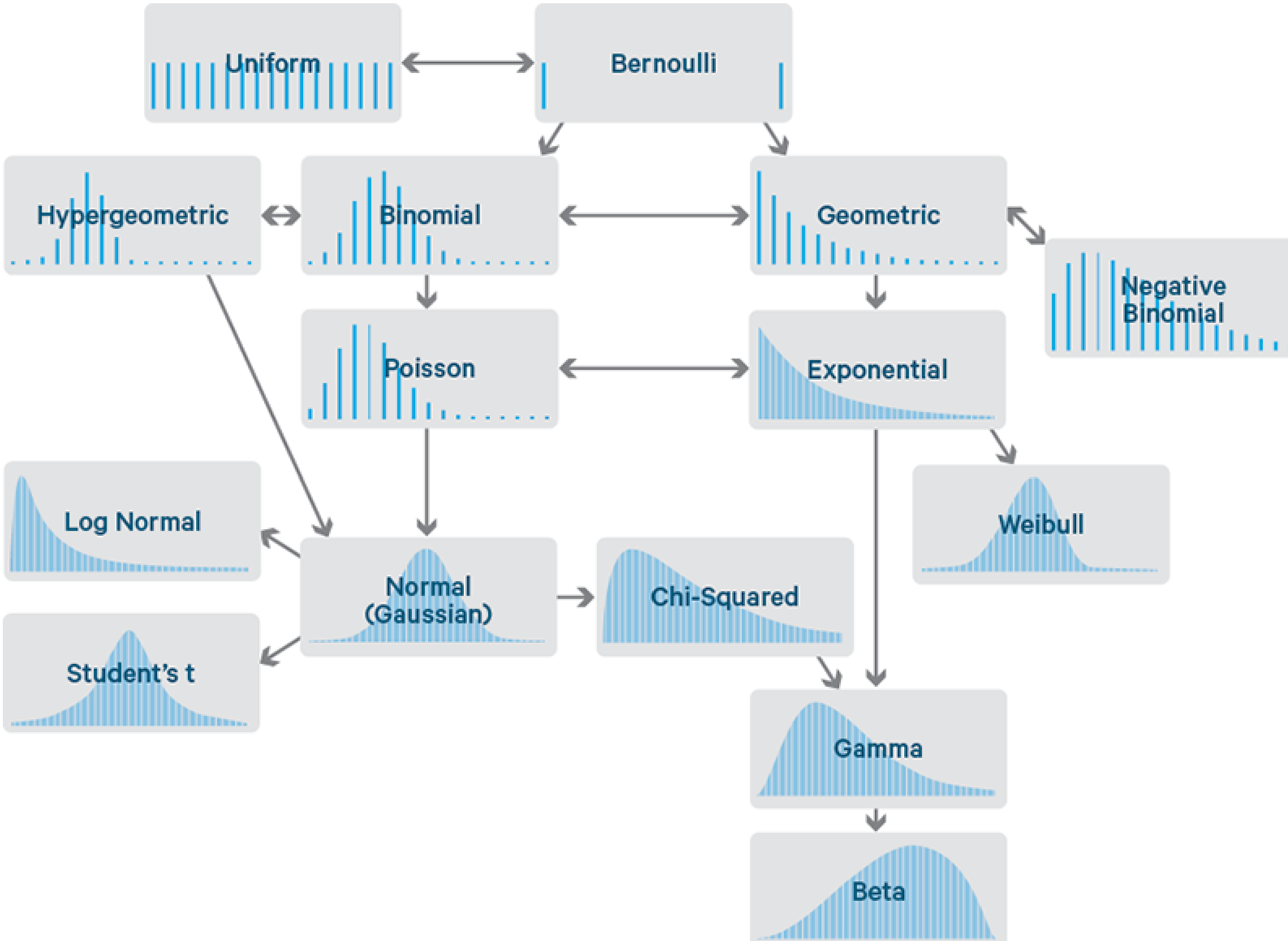
Lav R. Varshney

22 September 2017

Banach's smoking problem

- A smoker mathematician carries one matchbox in his right pocket and one in his left pocket. Each time he wants to light a cigarette, he selects a matchbox from either pocket with probability ½, independently of earlier selections. The two matchboxes have initially n matches each. What is the pmf of the number of remaining matches at the moment the mathematician reaches for a match and discovers the corresponding matchbox is empty?

Common probability distributions

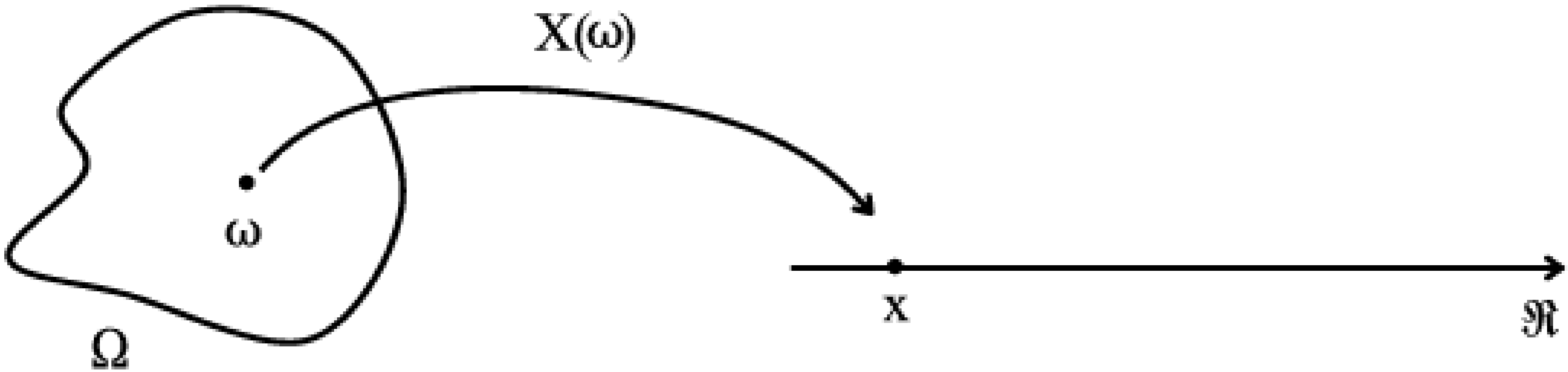


[https://blog.cloudera.com/blog/2015/12/common-probability-distributions-the-data-scientists-crib-sheet/]

Bernoulli process

- A Bernoulli process is an infinite sequence, $X_1, X_2, X_3, \ldots$ of Bernoulli random variables, all with the same parameter p, and independent of each other
- For any ω in the underlying probability space, the Bernoulli process has a corresponding realized value $X_k(\omega)$ for each time k, and that function of time is called the *sample path* of the Bernoulli process for outcome ω

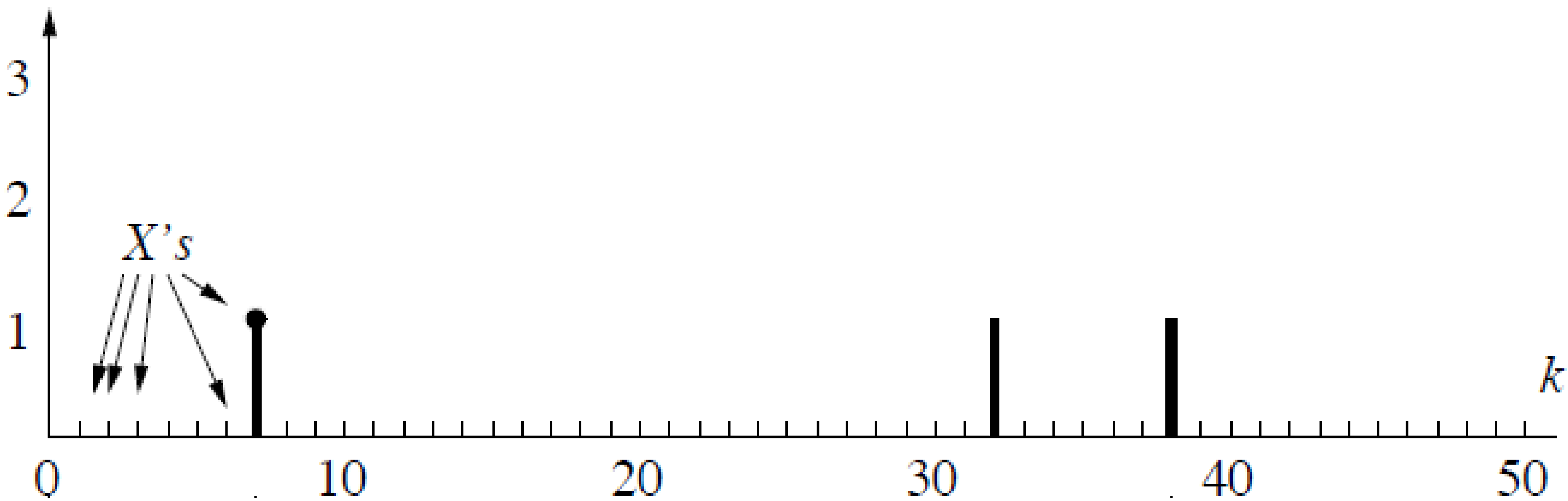
Random process is a mapping



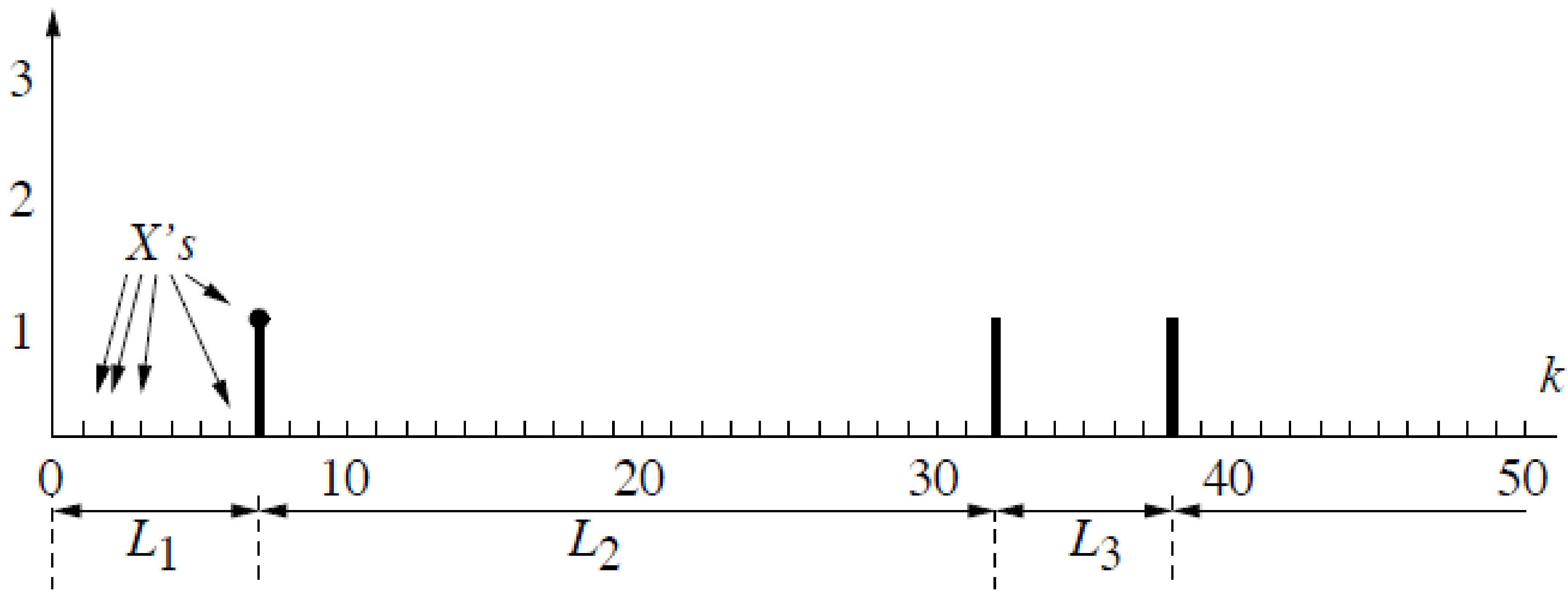
[cnx.org]

(recall the binary expansion of a real number)

Bernoulli process



Bernoulli process



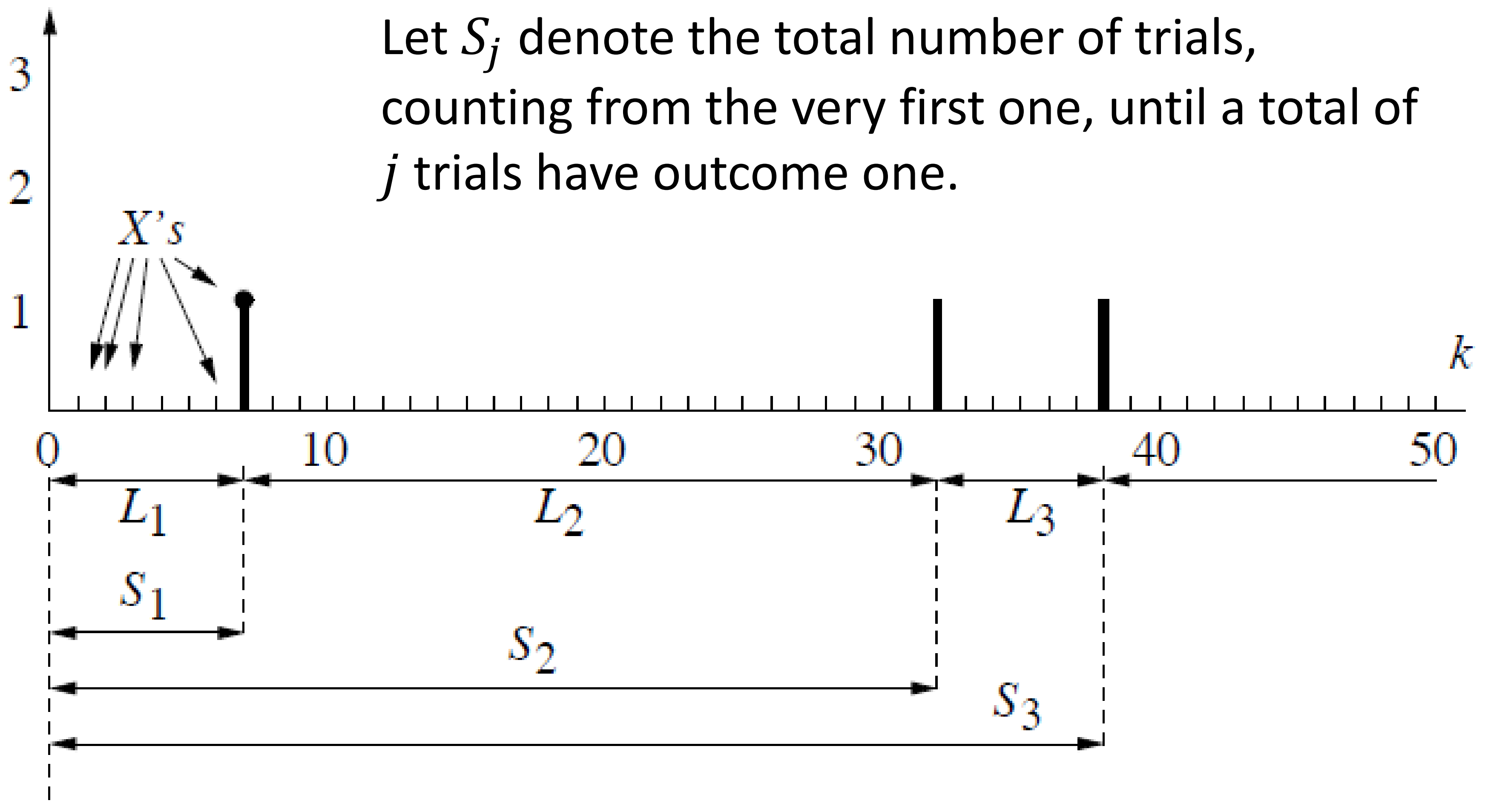
- Let L_1 be number of trials until outcome of a trial is one.
- Let L_2 be number of trials, after first L_1 trials, until outcome of a trial is one again….

Interevent interval

- What is the distribution of $L_1, L_2, \ldots$?

Bernoulli process

Let S_j denote the total number of trials, counting from the very first one, until a total of j trials have outcome one.



Cumulative numbers of trials for j ones

- Another way to express this is as $S_j = L_1 + L_2 + \cdots + L_j$ for $j \geq 1$
- What is the distribution of $S_1, S_2, \ldots$

Negative binomial random variable

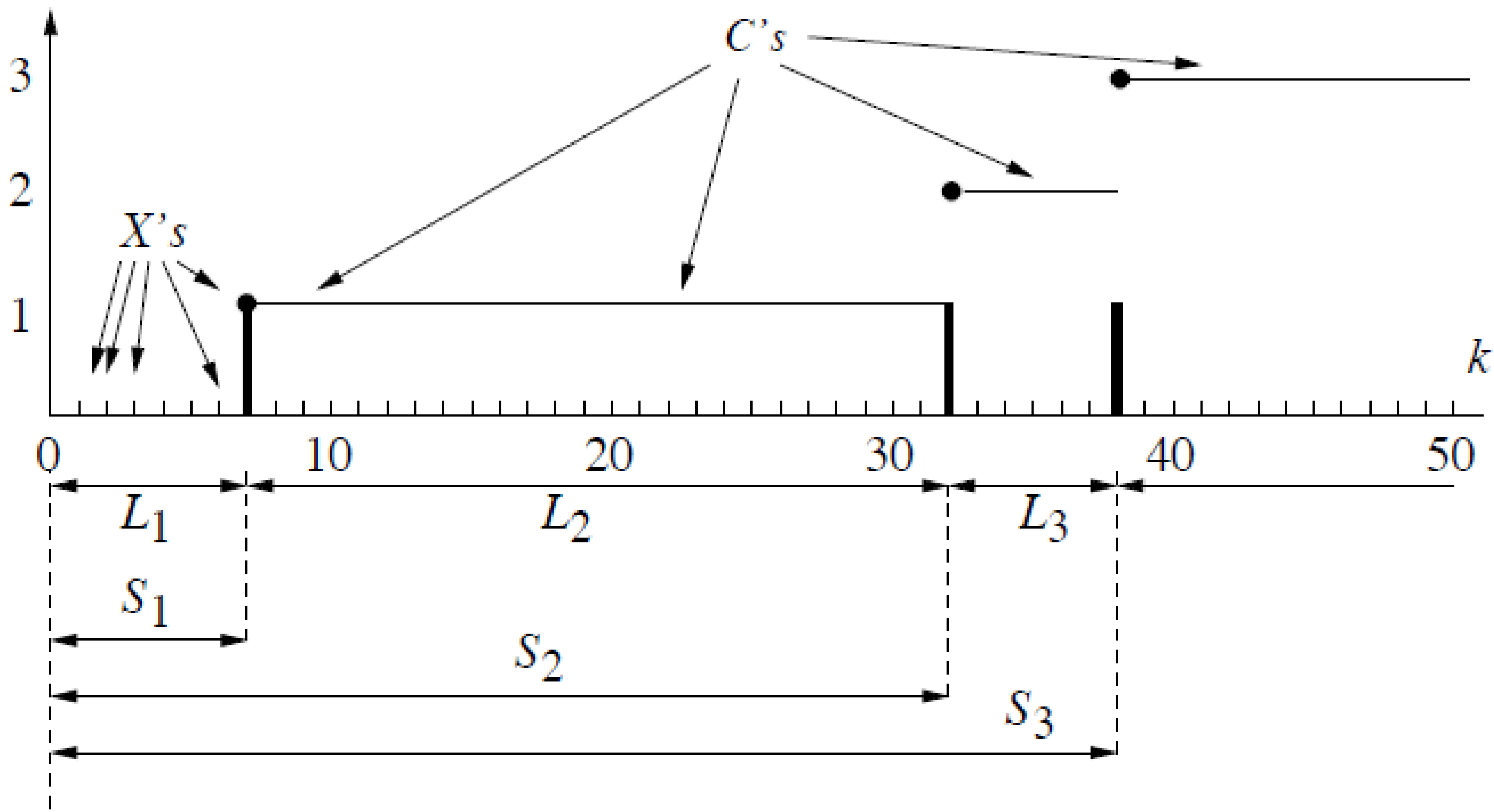
- The support for the pmf of S_r is $r, r+1, r+2, \ldots$
- So let $n \geq r$ and let $k = n - r$
- The event $\{S_r = n\}$ is determined by the outcomes of the first n trials.

Negative binomial distribution

- The event is true if and only if there are $r-1$ ones and k zeros in the first $k+r-1$ trials, and trial n is a one.
- There are $\binom{n-1}{r-1}$ such sequences of length n and each has probability $p^{r-1}(1-p)^{n-r}p$
- Pmf of S_r is given by

$$p(n) = \binom{n-1}{r-1} p^{r-1}(1-p)^{n-r}p, \qquad n \geq r$$

Bernoulli process



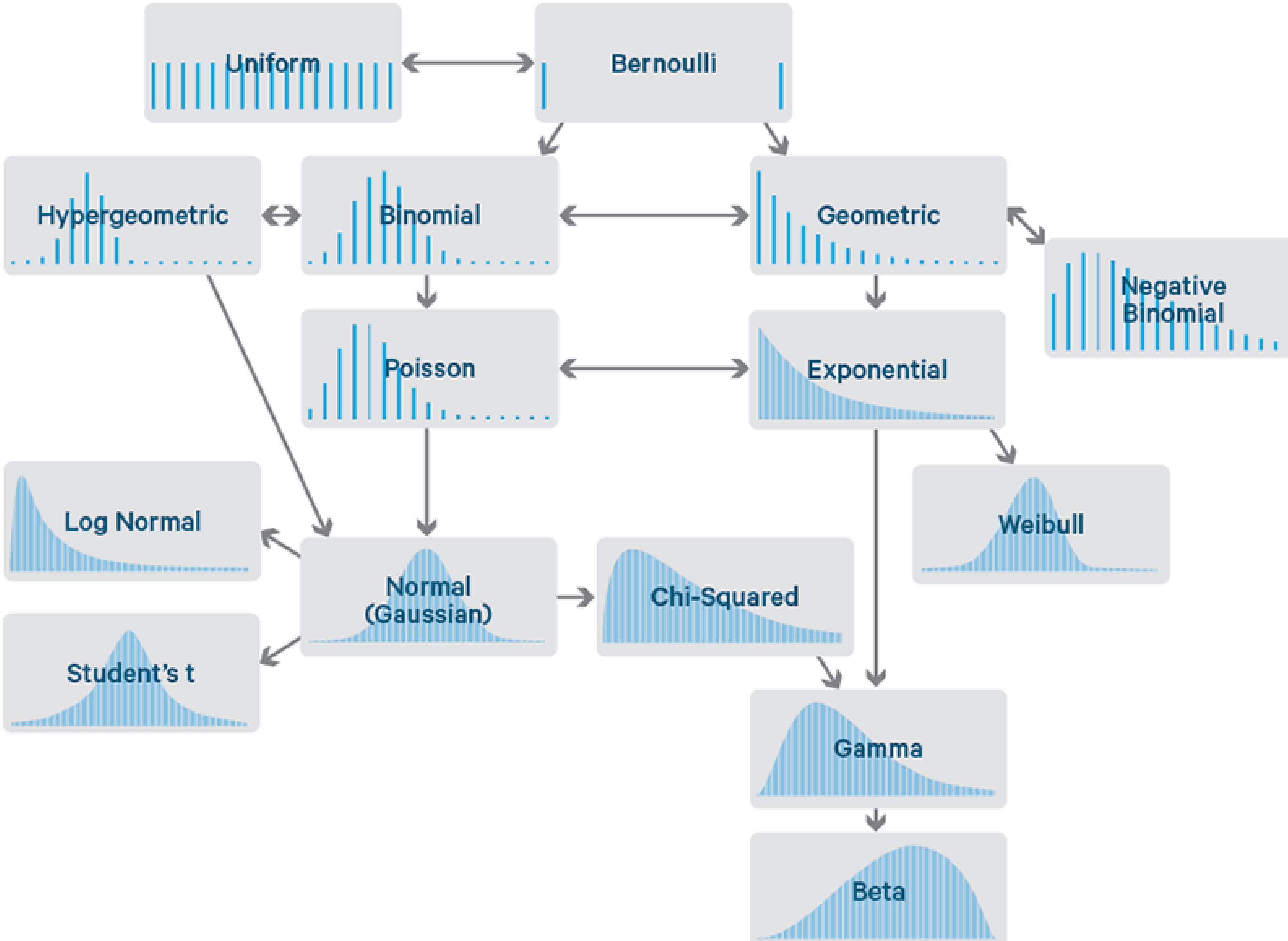
Bernoulli process

- Consider cumulative number of ones in k trials
- *Counting sequence* of the Bernoulli process
  - counts number of ones vs. number of trials
- The count $C_k = X_1 + X_2 + \cdots + X_k$
  - By convention, $C_0 = 0$
- What is the distribution?

Counting sequence

- For k fixed, C_k is the number of ones in k independent Bernoulli trials, so it has the binomial distribution with parameters k and p

Common probability distributions



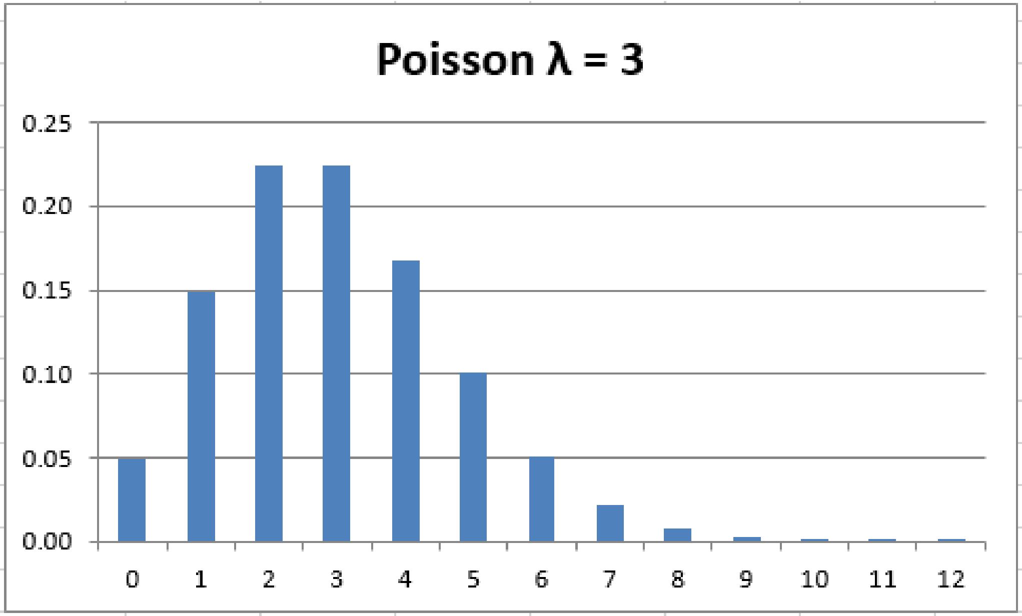
[https://blog.cloudera.com/blog/2015/12/common-probability-distributions-the-data-scientists-crib-sheet/]

Poisson distribution

- Poisson pmf with parameter $\lambda > 0$:

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

for $k \geq 0$



mean, variance both λ

[http://www.real-statistics.com/binomial-and-related-distributions/poisson-distribution/]

Poisson distribution

- The Poisson distribution arises frequently in practice, because it is a good approximation for a binomial distribution with parameters n and p, when n is very large, p is very small, and $\lambda = np$
- Rare-events limit
  - Radioactive emissions in fixed time interval
  - Incoming phone calls in fixed time interval
  - Misspelled words in document

Rare events

4. Beispiel: Die durch Schlag eines Pferdes im preufsischen Heere Getöteten.

(4. Example: Those killed in the Prussian army by a horse's kick.)

- L. J. Bortkiewicz, a Russian economist and statistician of Polish ancestry, published a book about the Poisson distribution, titled *The Law of Small Numbers*, where he made Prussian horse-kicking data famous. The data gave the number of soldiers killed by being kicked by a horse each year in each of 14 cavalry corps over a 20-year period.

Poisson distribution

Theorem If $\lim_{n\to\infty} np_n = \lambda > 0$, then

$$\lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Proof

You be MasterProbo