Special Random Variables

Motivations

- Some random variables occur frequently in data engineering applications
- They are also used as approximate models in analytical calculations, e.g. assume the message arrival time follows the Poisson distribution, assume the distribution is normal, etc.
- Below we study their properties. Then later we study how we can estimate their parameters

Bernoulli random variable

A Bernoulli random variable X with parameter p describes the outcome of an experiment which has two outcomes: "success" or "failure". It can be used as an indicator variable. It is named after the mathematician James Bernoulli.

Example: Throw a coin and obtain "head"

Probability mass function

$$P\{X = 1\} = p \qquad\qquad P\{X = 0\} = 1 - p$$

$0 \leq p \leq 1$ is the probability that the outcome is a success

Expected value

$$E[X] = 1 \cdot P\{X = 1\} + 0 \cdot P\{X = 0\} = p$$

Variance

$$Var[X] = E[X^2] - (E[X])^2$$

$$E[X^2] = (1)^2 \cdot P\{X = 1\} + (0)^2 P\{X = 0\} = p$$

Hence

$$Var[X] = p - p^2 = p(1 - p)$$

Binomial random variable

A binomial random variable X with parameters (n, p) is the sum of n independent Bernoulli random variables X_i, $i = 1, \ldots, n$, i.e.,

$$X = \sum_{i=1}^{n} X_i$$

Example: Throw a coin 8 times and the number of heads is a binomial random variable

Probability mass function

$$P\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \ldots, n$$

$\binom{n}{i}$ is the number of ways that the i successes can occur

$p^i(1-p)^{n-i}$ is the probability that a particular way occurs

Example: There are $(8)(7)/2$ ways that 2 "heads" can occur in 8 coin throws. In a biased coin with $p = 0.6$, each way has a probability of $(0.6)^2(0.4)^6$ of occurring

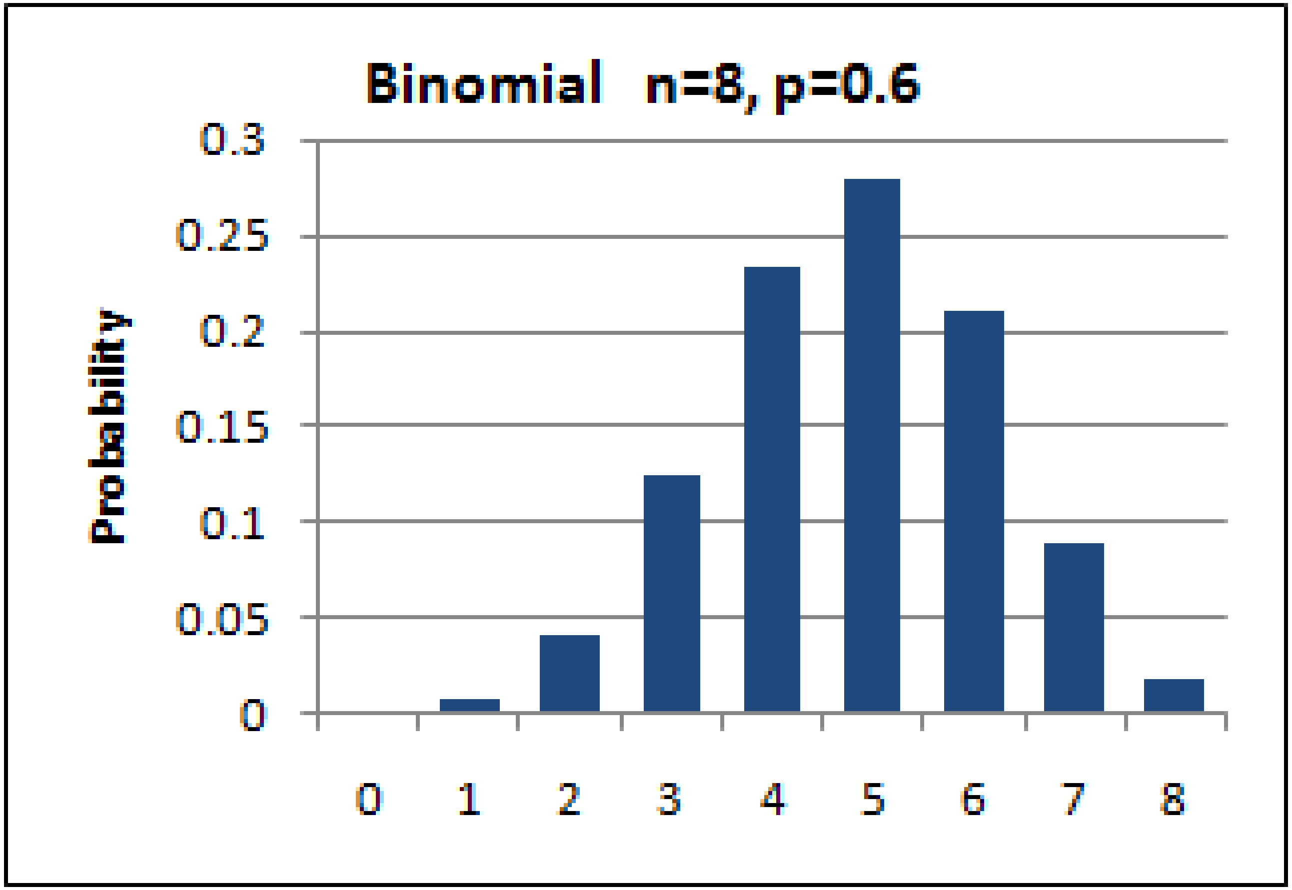
Expected value

$$E[X] = \sum_{i=1}^{n} E[X_i] = np$$

Variance

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = np(1 - p)$$

An example of the binomial distribution is as shown. It is a discrete probability distribution



Given a biased coin with probability 0.6, the number of “heads” in 8 throws. Observe the shape and check whether it agrees with your intuition

Poisson random variable

A Poisson random variable X with parameter parameter λ, $\lambda > 0$ can take on values $0, 1, 2, \ldots$,

Its probability mass function is given by

$$P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!} \qquad\qquad i = 0,1,\ldots$$

It is named after the mathematician S.D. Poisson.

The expected value and variance can be derived using moment generating function

$$\phi(t) = E[e^{tX}]$$
$$= \sum_{i=0}^{\infty} e^{ti}\, e^{-\lambda} \frac{\lambda^i}{i!}$$
$$= e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!}$$

Using the exponential theorem $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$

$$\phi(t) = e^{-\lambda} e^{\lambda e^t} = exp\{\lambda(e^t - 1)\}$$

where $exp\{\quad\}$ is another way to write $e^{\{\quad\}}$

Take the 1st and 2nd derivatives and then set $t = 0$

$$\phi'(t) = \lambda e^t \exp\{\lambda(e^t - 1)\}$$
$$\phi''(t) = (\lambda e^t)^2 \exp\{\lambda(e^t - 1)\} + \lambda e^t \exp\{\lambda(e^t - 1)\}$$

Expected value

$$E[X] = \phi'(0) = \lambda$$

Variance

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \\ &= \phi''(0) - (E[X])^2 \\ &= (\lambda^2 + \lambda) - (\lambda)^2 = \lambda \end{aligned}$$

Note that it has the interesting property that both the mean and the variance is equal to λ

Usage

It is used as an approximation for a binomial random variable with parameters (n, p) when

n is large and p is small

Physical meaning: calculate the number of occurrences of some events for which the success probability is small but we have many attempts

In this case, $\lambda = np$

Examples

1. The number of misprints on the text book
2. The number of Hong Kong citizens living to 100 years of age
3. The number of wrongly dialed telephone number in a year
4. The number of mobile phones that fail on their first month of use

Derivations of the approximation

Starting with the binomial distribution and let $\lambda = np$

$$P\{X = i\} = \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}$$

$$= \frac{n!}{i!(n-i)!} \left(\frac{\lambda}{n}\right)^i (1 - \frac{\lambda}{n})^{n-i}$$

$$= \frac{n(n-1)\ldots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

For large n and small p,

$$(1-\lambda/n)^n \approx e^{-\lambda}$$

[because by the exponential theorem

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Put in $x = -p$

$$e^{-p} = 1 - \frac{p}{1!} + \frac{p^2}{2!} - \frac{p^3}{3!} + \cdots$$

When p is small,

$$e^{-p} \approx 1 - p$$

Thus

$$(1 - \lambda/n)^n = (1 - p)^n \approx e^{-pn} = e^{-\lambda}]$$

$$\frac{n(n-1)\ldots(n-i+1)}{n^i} \approx 1$$

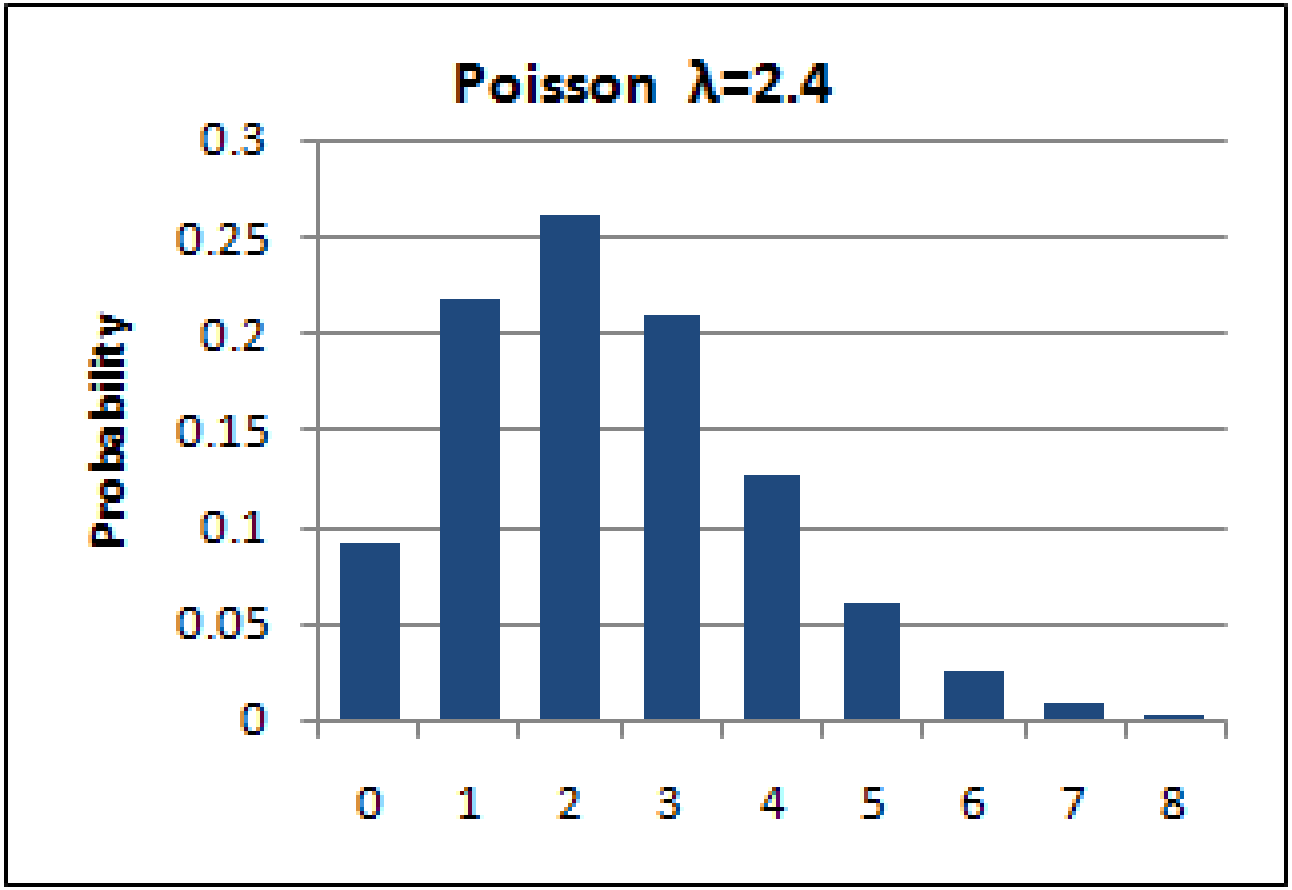
$$\left(1-\frac{\lambda}{n}\right)^i \approx 1$$

Thus large n and small p,

$$P\{X=i\} = e^{-\lambda}\frac{\lambda^i}{i!}$$

which is the Poisson distribution.

An example of the Poisson distribution is as shown. It is a discrete probability distribution



Example: The number of wrong spellings in a paper with 2400 words and $p = 0.001$. Type one wrong word in 1000 words. Does the distribution agree with your intuition?

Geometric random variable

Suppose independent trials, each having a probability p, $0 < p < 1$, of being a success, are performed. A geometric random variable X is the number of trials until the first success occur.

Example: The number of attempts until passing a test

Its probability mass function is given by

$$P\{X = n\} = (1 - p)^{n-1}p \qquad n = 1,2, \ldots$$

This is the probability that the first $(n - 1)$ trial fails and the last trial succeeds.

Expected value

Let $q = 1 - p$

$$E[\mathrm{X}] = \sum_{i=1}^{\infty} iq^{i-1}p = \sum_{i=1}^{\infty}(i - 1 + 1)q^{i-1}p$$

$$= \sum_{i=1}^{\infty}(i - 1)q^{i-1}p + \sum_{i=1}^{\infty} q^{i-1}p$$

$$= \sum_{j=0}^{\infty} jq^{j}p + 1$$

$$= q\sum_{j=1}^{\infty} jq^{j-1}p + 1 = qE[X] + 1$$

Hence

$$E[X] = \frac{1}{p}$$

The expected number of trials until the first success is $\frac{1}{p}$

Example: The expected number of dice throw to get the first “6” is 6

Variance

$$E[X^2] = \sum_{i=1}^{\infty} i^2 q^{i-1} p = \sum_{i=1}^{\infty} (i - 1 + 1)^2 q^{i-1} p$$

$$= \sum_{i=1}^{\infty}(i-1)^2 q^{i-1}p + \sum_{i=1}^{\infty} 2(i-1)q^{i-1}p + \sum_{i=1}^{\infty} q^{i-1}p$$

$$= \sum_{j=0}^{\infty} j^2 q^j p + 2\sum_{j=0}^{\infty} j q^j p + 1$$

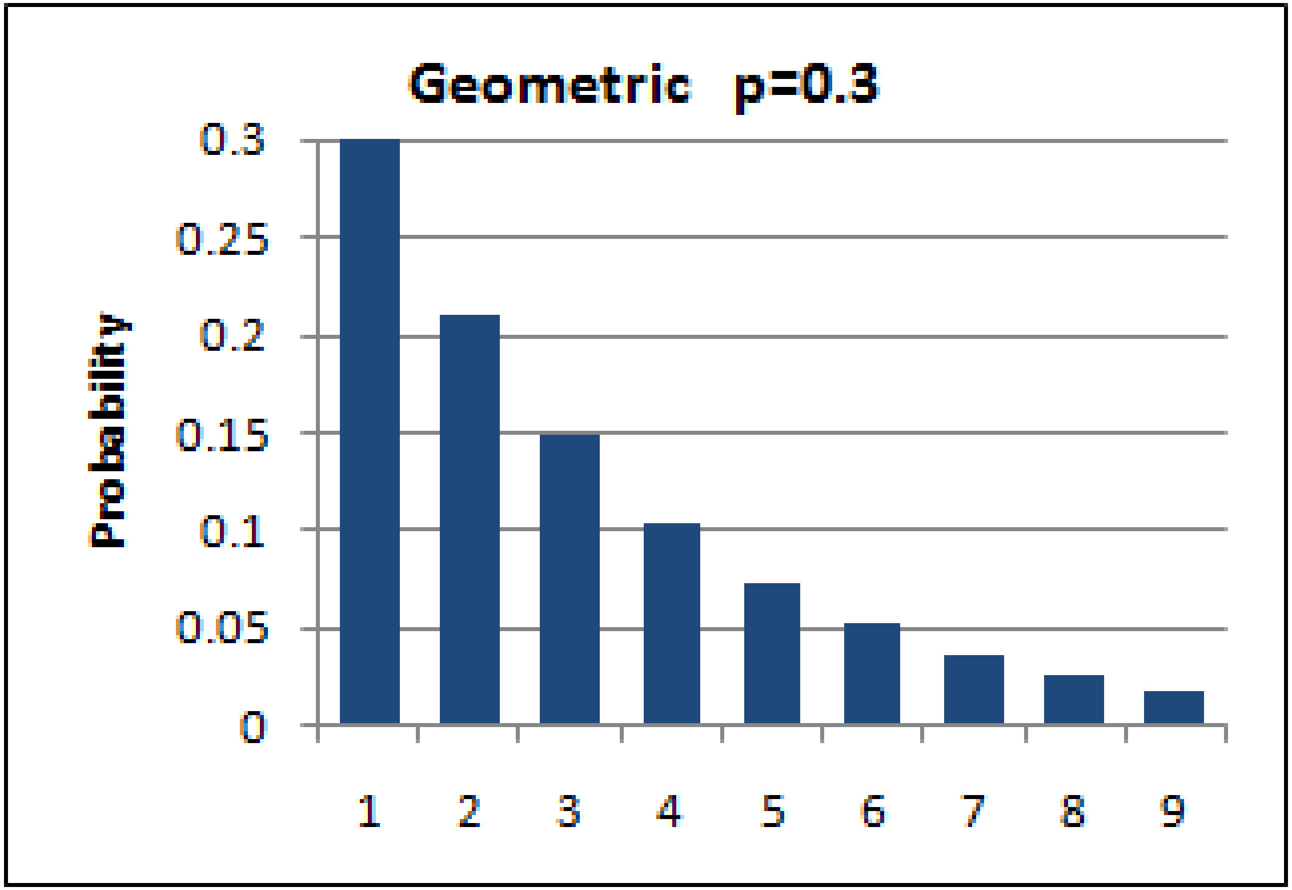
$$= qE[X^2] + 2qE[X] + 1$$

This gives

$$pE[X^2] = \frac{2q}{p} + 1 \quad \text{or } E[X^2] = \frac{2q+p}{p^2} = \frac{q+1}{p^2}$$

$$Var[X] = E[X^2] - E[X^2] = \frac{1-p}{p^2}$$

An example of the geometric distribution is shown. It is a discrete probability distribution



Example: Number of attempts before passing a test

Exponential random variable

A continuous random variable X is an exponential random variable with parameter λ if its probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases}$$

The exponential distribution often arises, in practice, as being the distribution of the amount of time until some specific events occurs

A non-negative random variable X is memoryless if

$$P\{X > s + t | X > t) = P\{X > s\} \qquad for\ all\ s, t \geq 0$$

The exponential distribution is memoryless.

In fact, it can be shown that the exponential distribution is the only continuous random variable which is memoryless

Example

1. The time it takes for a radioactive particle to decay
2. The time it takes for the next phone call to be a wrong number
3. The time it takes for the next phone call (?)

Using the moment generating function

$$\phi(t) = E[e^{tX}]$$
$$= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$
$$= \lambda \int_0^\infty e^{-(\lambda - t)x}\, dx$$
$$= \frac{\lambda}{\lambda - t} \qquad\qquad t < \lambda$$

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2}$$
$$\phi''(t) = \frac{2\lambda}{(\lambda - t)^3}$$

Expected value

$$E[X] = \phi'(0) = 1/\lambda$$

Variance

$$\begin{aligned} Var(X) &= \phi''(0) - (E[X])^2 \\ &= 2/\lambda^2 - 1/\lambda^2 \\ &= 1/\lambda^2 \end{aligned}$$

Memoryless property

A non-negative random variable X is memoryless if

$$P\{X > s + t | X > t) = P\{X > s\} \qquad for\ all\ s, t \geq 0$$

This is equivalent to

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

or

$$P\{X > s + t\} = P\{X > s\}P\{X > t\}$$

When X is an exponential random variable,

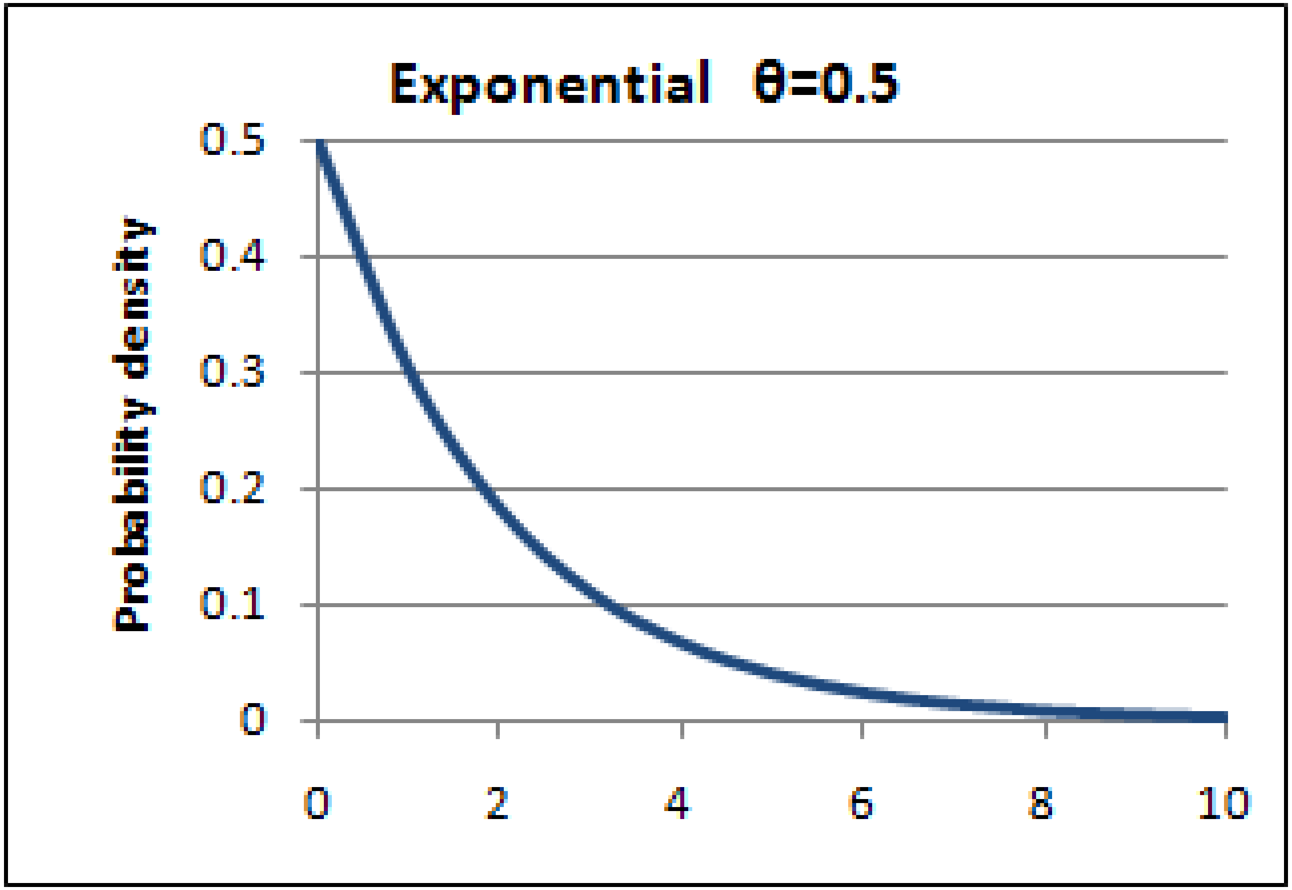
$$F(x) = P\{X \leq x\}$$
$$= \int_0^x \lambda e^{-\lambda y} dy$$
$$= 1 - e^{-\lambda x} \qquad x \geq 0$$

Thus

$$P\{X > x\} = e^{-\lambda x} \qquad x > 0$$

As $e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}$, the exponential random variable is memoryless.

An example of the exponential distribution is as shown. It is a continuous probability distribution



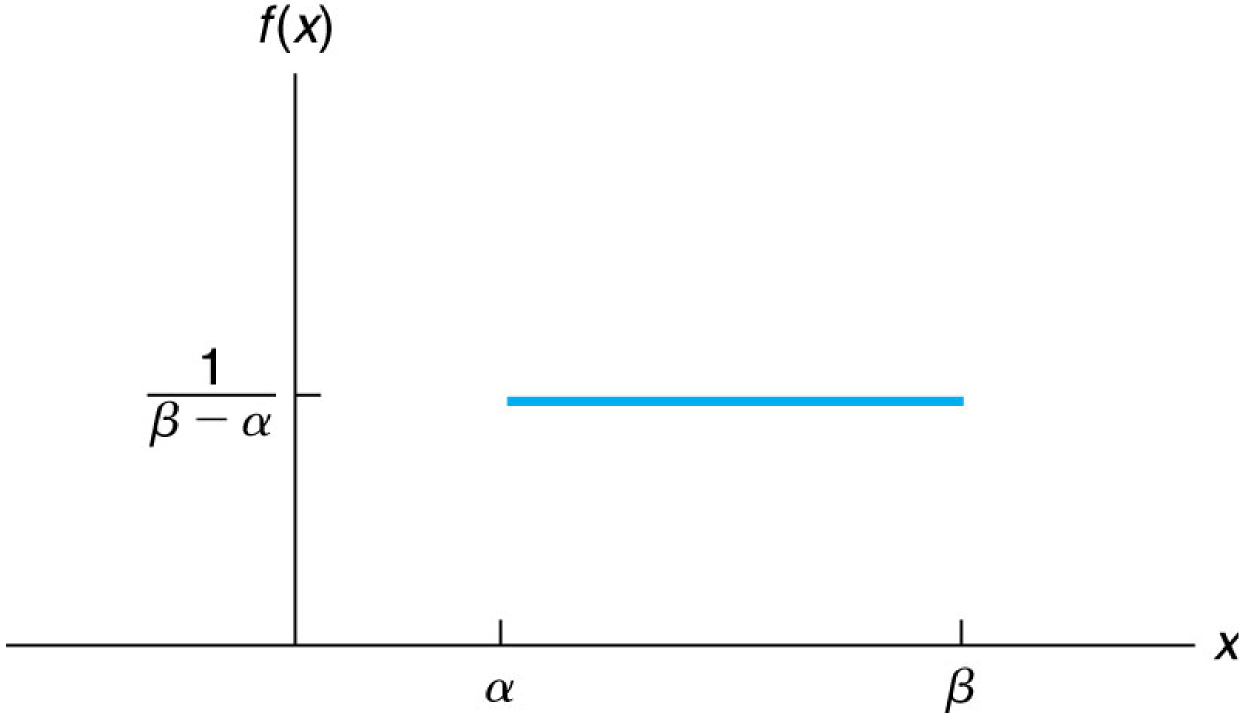
Uniform Distribution

- A continuous random variable X is said to be uniformly distributed over the interval $[\alpha, \beta]$ if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & if\ \alpha \le x \le \beta \\ 0 & otherwise \end{cases}$$

- The uniform distribution arises in practice when we suppose a certain random variable is equally likely to be near any value in the interval $[\alpha, \beta]$

The uniform distribution is a continuous probability distribution



Example: Calculator generating a pseudo-random number between $\alpha = 0$ and $\beta = 1$

Other Commonly Used Non-normal Distributions

- Discrete: negative binomial, hypergeometric, zeta(or zipf)
- Continuous: gamma, Weibull, Cauchy, Beta
- For more information, consult the reference A First Course in Probability
- Below we briefly introduce some. You may encounter them later in your career

Negative binomial random variable

Suppose that independent trials, each having probability p, $0 < p < 1$, of being a success are performed until a total of r successes is accumulated. If we let X equal the number of trials required, then

$$P\{X = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r} \qquad n = r, r+1, \ldots$$

Because in order for the rth success to occur at the nth trial, there must be $r - 1$ successes in the first $n - 1$ trials, the probability of this event is

$$\binom{n-1}{r-1} p^{r-1} (1-p)^{n-r}$$

The last trial must be a success, which occur with a probability p

X is called a negative binomial random variable with parameter (r, p)

A geometric random variable corresponds to the special case of $(1, p)$

Hypergeometric random variable

A bin contains $N + M$ batteries, of which N are of acceptable quality and the other M are defective. A sample of size n is to be randomly chosen (without replacement) in the sense that the set of sampled batteries is equally likely to be any of the $\binom{N+M}{n}$ subsets of size n. Let X denote the number of acceptable batteries in the sample, then

$$P\{X = i\} = \frac{\binom{N}{i}\binom{M}{n-i}}{\binom{N+M}{n}} \qquad i = 0, 1, \ldots, \min(N, n)$$

Weibull random varaible

A random variable X whose probability density function is of the form

$$f(x) = \begin{cases} 0 & x \le v \\ \dfrac{\beta}{\alpha}\left(\dfrac{x-v}{\alpha}\right)^{\beta-1} exp\left\{-\left(\dfrac{x-v}{\alpha}\right)^{\beta}\right\} & x > v \end{cases}$$

It is widely used, in the field of life phenomena, as the distribution of the lifetime of some object, particularly when the “weakest link” model is appropriate for the object (the object has many parts and the object fails when any of its parts fail)

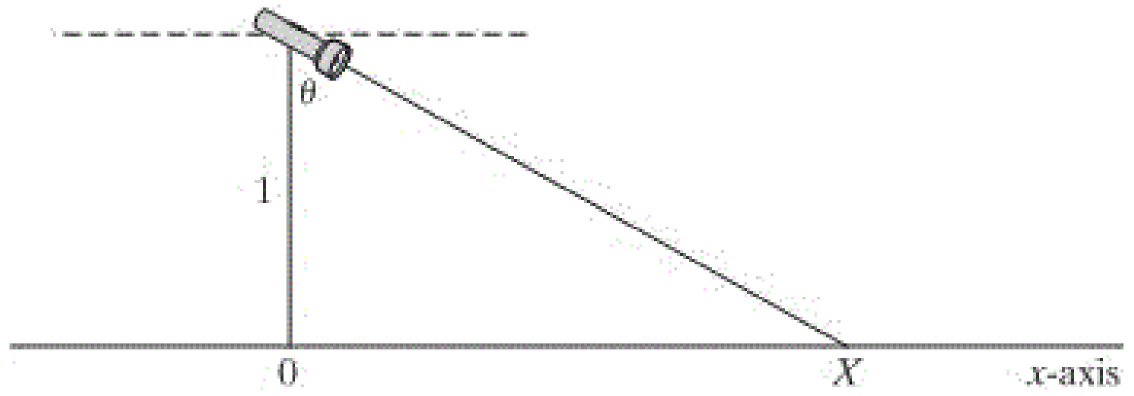
Cauchy random variable

A random variable h, as a Cauchy distribution with parameter α, $-\infty < \alpha < \infty$, if its density is given by

$$f(x) = \frac{1}{\pi}\frac{1}{1+(x-\alpha)^2} \qquad -\infty < x < \infty$$

Example of Cauchy distribution

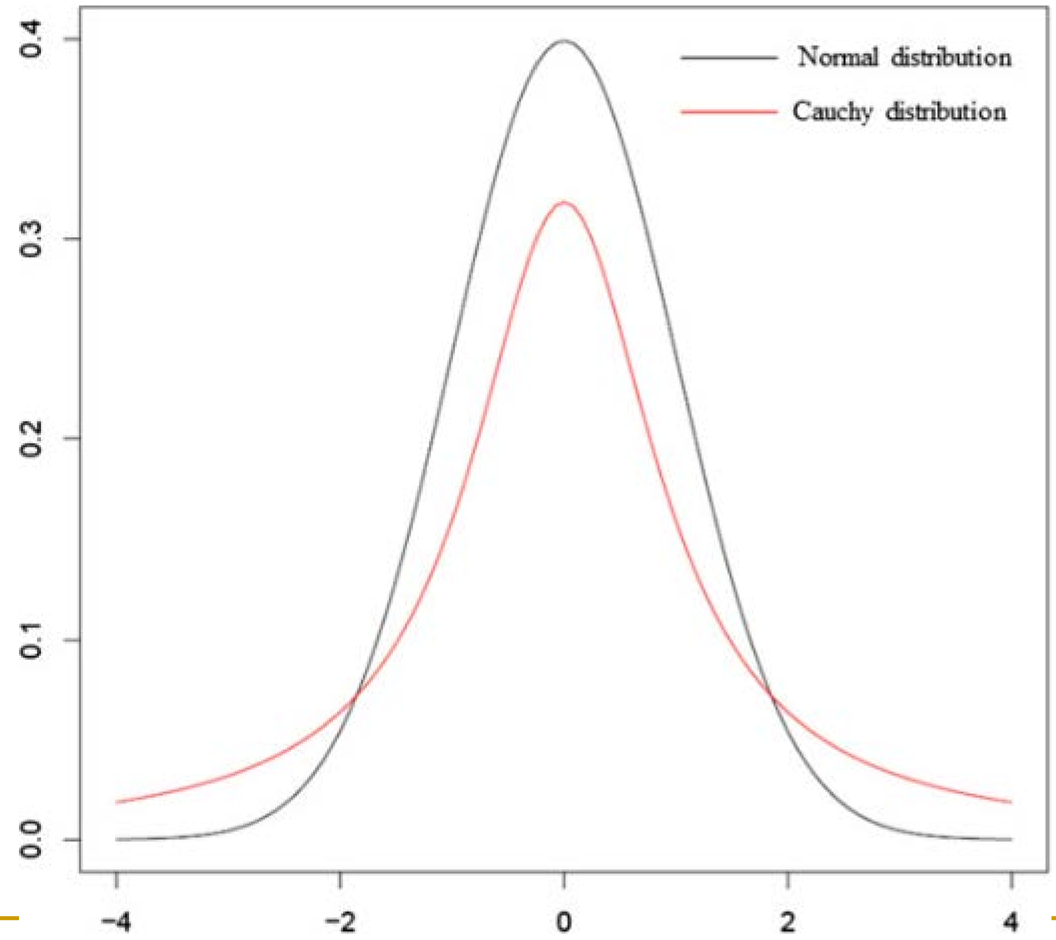
Flashlight spins randomly with an angle of $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ and stops



The random variable X has the Cauchy distribution

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

The Cauchy distribution has a shape like the normal distribution but has a heavier tail. That is, rare events are more likely to happen for Cauchy distribution



References

- Text Ch. 5-7
- Geometric distribution: See A First Course in Probability pg. 158-160
- Non-normal distributions
  - A First Course In Probability
  - http://www.icse.xyz/msor/psme/standard.html teaches how to use Excel to plot the probability distributions