UNIVERSITY OF LINCOLN

SCHOOL OF MATHEMATICS AND PHYSICS

MTH 1005
PROBABILITY AND STATISTICS

Semester B
Lecture 9
(14/3/2023)

Danilo Roccatano

Office: INB 3323
Email: droccatano@lincoln.ac.uk

SPECIAL DISTRIBUTIONS

Learning objectives:

Define and use the properties of the major probability distributions:

- **Uniform**
- Binomial
- Geometric
- Poisson
- Exponential
- Normal

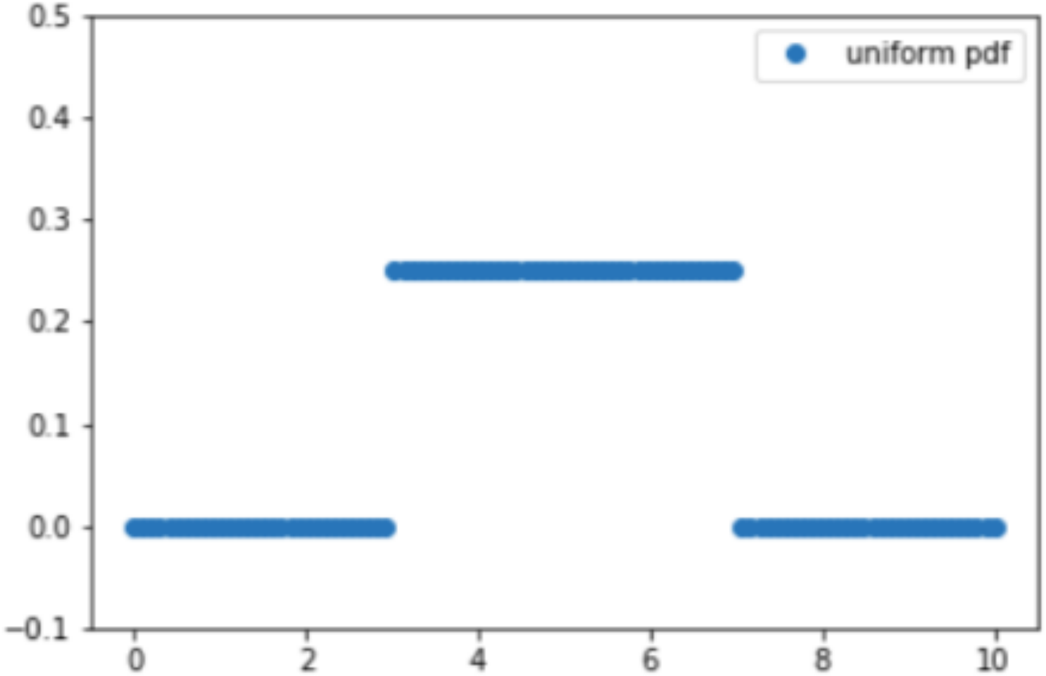
with examples where these distributions occur.

SUMMARY OF THE UNIFORM DISTRIBUTION

A continuous random variable X is said to be uniformly distributed over the interval [a, b] if its probability is given

$$P(\alpha \leq x \leq \beta) = \frac{\beta - \alpha}{b - a} \qquad \text{PMF}$$

Then it is a **uniform distribution** with



EXPECTATION VALUE

The expectation value of a uniformly distributed random variable is calculated as

$$
\begin{aligned}
\mathrm{E}[X] &= \int_a^b x f(x) \mathrm{d}x \\
&= \frac{1}{(b-a)} \int_a^b x \mathrm{d}x \\
&= \frac{b^2 - a^2}{2(b-a)} \\
&= \frac{(b-a)(b+a)}{2(b-a)} \\
&= \frac{a+b}{2}
\end{aligned}
$$

So the expected value of the random variable is the midpoint of the interval that it is de ned over.

VARIANCE OF UNIFORMLY DISTRIBUTED RANDOM VARIABLE

We use our second definition of variance

$$\sigma_X^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

We have already found E[X] , we need to calculate E[X^2].

$$\begin{aligned}
\mathrm{E}[X^2] &= \int_a^b x^2 f(x)\mathrm{d}x \\
&= \frac{1}{(b-a)} \int_a^b x^2 \mathrm{d}x \\
&= \frac{b^3 - a^3}{3(b-a)} \\
&= \frac{b^2 + ab + a^2}{3}
\end{aligned}$$

VARIANCE OF UNIFORMLY DISTRIBUTED RANDOM VARIABLE

and so we get

$$\begin{aligned}\sigma_X^2 &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\ &= \frac{b^2+ab+a^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{(b-a)^2}{12}\end{aligned}$$

QUANTILES

We can find the quantiles of the uniform distribution very simply.

The cumulative probability distribution function of a uniformly distributed continuous random variable defined between a and b is given by a

$$
\begin{aligned}
F(\alpha) &= \int_{-\infty}^{a} f(x)\mathrm{d}x \\
&= \frac{1}{(b-a)} \int_{a}^{\alpha} \mathrm{d}x \\
&= \frac{(\alpha - a)}{(b-a)}
\end{aligned}
$$

this is just the ratio of the length of the interval the variable is de ned over to that of lower bound of the interval to the value of the variable at which we want to calculate the cumulative distribution function.

Example

Find the value of p_{95} such that there is a 95% probability that the random variable X takes a value smaller than p_{95}.

$$
\begin{aligned}
P(X < p_{95}) = F(p_{95}) &= 0.95 \\
&= \frac{(p_{0.95} - a)}{(b-a)} \\
\implies p_{0.95} &= 0.95(b-a) + a
\end{aligned}
$$

also implying the median is at the centre of the distribution, like the mean.

RANDOM NUMBER

Probably, the most important use of the uniform distribution is in defining a random number.

The value of a uniform [0,1] random variable is called a random number.

We used computer approximations to this distribution in our simulations.

In Scientific Computing module, you will more about the generation of random numbers using computers.

SPECIAL DISTRIBUTIONS

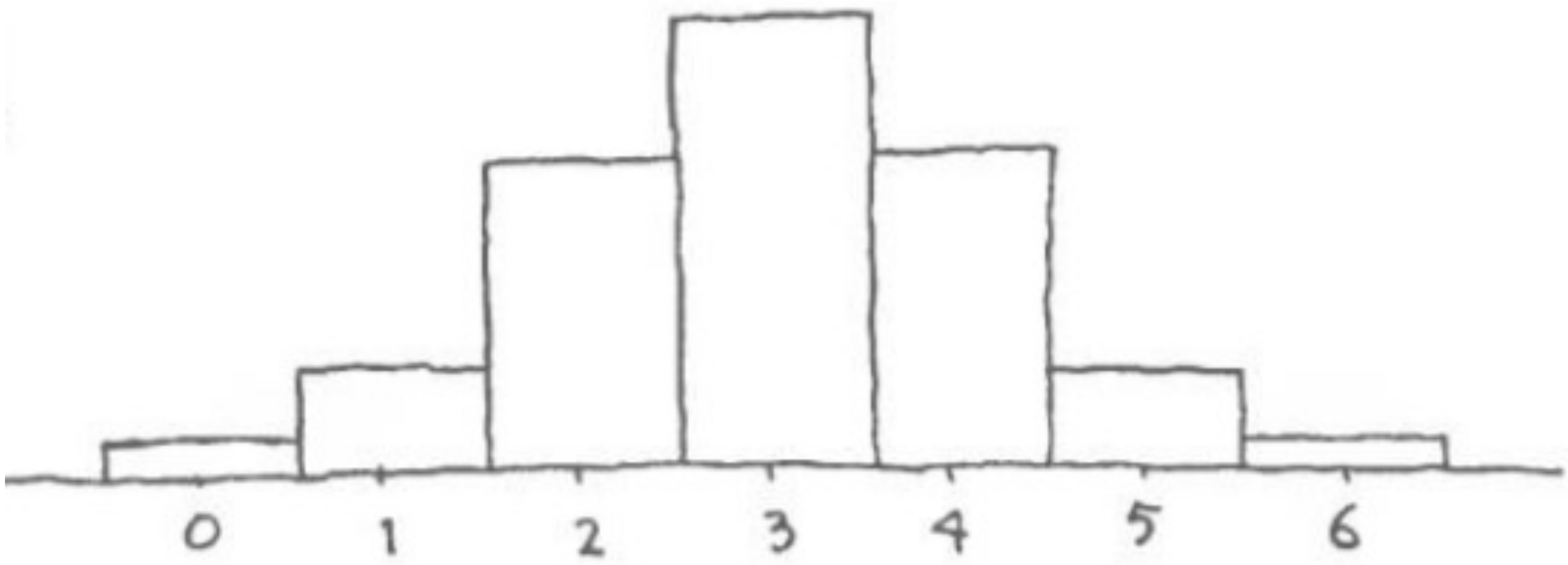
Learning objectives:

Define and use the properties of the major probability distributions:

- Uniform
- **Binomial**
- Poisson
- Geometric
- Exponential
- Normal

with examples where these distributions occur.

BINOMIAL DISTRIBUTION



BINOMIAL DISTRIBUTION

We've used this fairly often in the first few weeks.

Definition

It is the discrete distribution that arises when we have an experiment that can be partitioned into two mutually exclusive events ({S} , success and {F } , failure) with a probability of success P ({S}) that is known, repeated independently n times.

Definition

A discrete random variable X with a probability mass function of the form:

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

is said to have a binomial distribution.

Where n is the number of repetitions of the experiment and p is the probability of 'success'.

CHECK IT IS A PROBABILITY MASS DISTRIBUTION

To be a probability mass distribution we must have

$$\sum\nolimits_{all\ x} P(x) = 1$$

For the binomial distribution we have

$$\begin{aligned}\sum_{\text{all } x} P(x) &= \sum_{\text{all } x} \binom{n}{x} p^x (1-p)^{n-x} \\ &= (p + (1-p))^n \\ &= 1^n = 1\end{aligned}$$

we use the binomial expansion

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

So it fits that criteria, and all *p(x)* are between 0 and 1.

THE MEAN AND THE VARIANCE

We have that the mean (expectation value) of our binomially distributed variable X is

$$
\begin{aligned}
\mathrm{E}[X] &= \sum_{\text{all } x} x p(x) \\
&= \sum_{\text{all } x} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n! p^x (1-p)^{n-x}}{(n-x)!x!} \quad (x = 0 \text{ term is zero}) \\
&= np \sum_{x=1}^{n} x \frac{(n-1)! p^{x-1} (1-p)^{n-x}}{(n-x)!x!} \quad (\text{factorize out } np)
\end{aligned}
$$

THE MEAN AND THE VARIANCE

- $\frac{x}{x!} = \frac{1}{(x-1)!}$
- $(n-x)!$ can be rewritten as $(n-1-(x-1))!$
- $(1-p)^{n-x}$ can be written as $(1-p)^{n-1-(x-1)}$ we then get

$$\mathrm{E}[X] = np\sum_{x=1}^{n}\frac{(n-1)!p^{x-1}(1-p)^{n-1-(x-1)}}{(n-1-(x-1))!(x-1)!}$$

and, we substitute y= x – 1 and N= n – 1 to get

$$\mathrm{E}[X] = np\sum_{y=0}^{N}\frac{N!p^{y}(1-p)^{N-y}}{(N-y)!y!}$$

$$= np[\sum_{y=0}^{N}\binom{N}{y}p^{y}(1-p)^{N-y}]$$

$$= np$$

THE MEAN AND THE VARIANCE

The variance can be found in a similar manner, but I won't give it all here.

$$\begin{aligned}\sigma_X^2 &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\ &= n^2p^2 - np^2 + np - n^2p^2 \\ &= np(1-p)\end{aligned}$$

Note: the same proof can be obtained in a more elegant way using moment generating functions.

SUMMARY OF EXPECTATION AND VARIANCE OF BINOMIAL DISTRIBUTION

EXPECTATION

$$\mathrm{E}[X] = np$$

VARIANCE

$$\sigma_X^2 = E[X^2] - E[X]^2 = np(1-p)$$

CUMULATIVE DISTRIBUTION FUNCTION

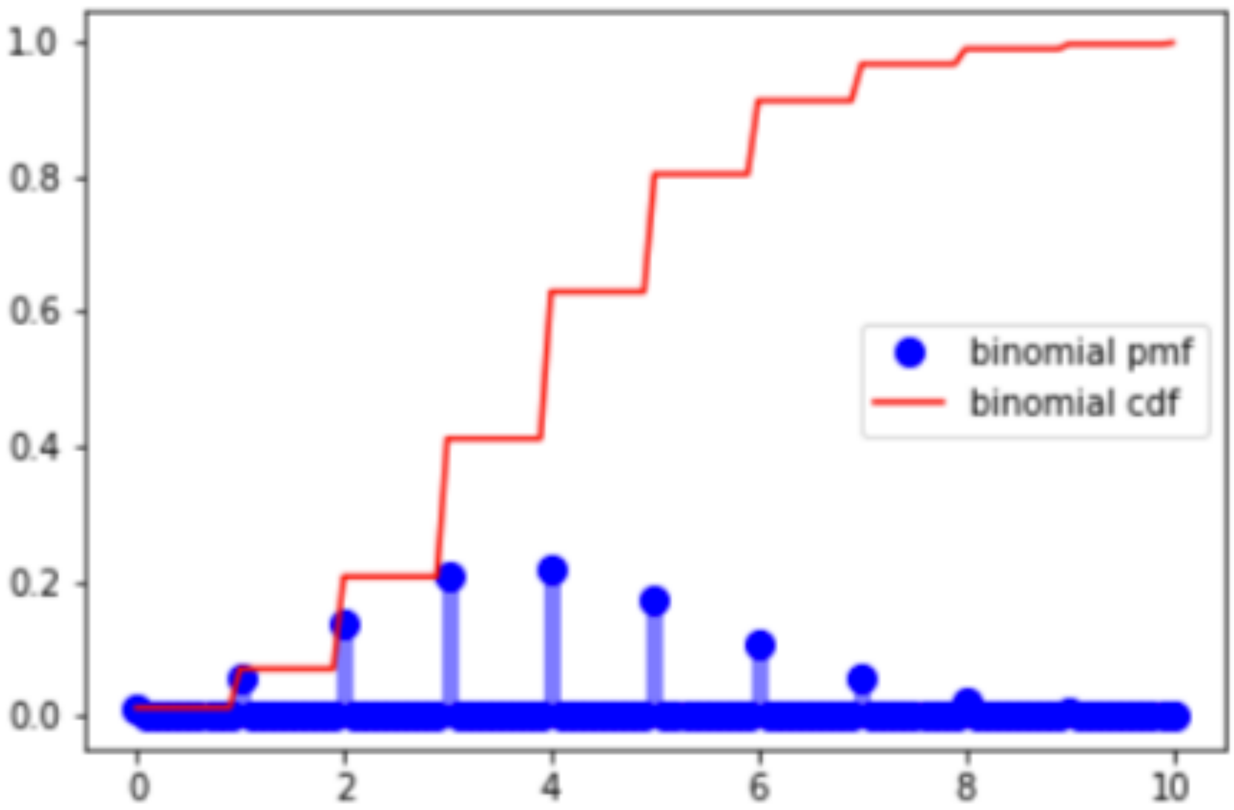
For discrete random variables, the cumulative distribution function is often available in tabulated form.

To calculate it we need to sum up all the *p(x)* values where x is less than the value of interest.

$$F(y) = \sum_{x \leq y} p(x)$$

this can get quite tricky for the binomial distribution.

CUMULATIVE DISTRIBUTION FUNCTION



CUMULATIVE DISTRIBUTION FUNCTION

we can access the quantiles for the binomial distribution like so

```
stats.binom.ppf(0.99, n, p)
6.0
```

This says that for our distribution (binomial with parameters n and p) a random value taken from this distribution will have a value less than than 6.99% of the time.

SPECIAL DISTRIBUTIONS

Learning objectives:

Define and use the properties of the major probability distributions:

- Uniform
- Binomial

- **Poisson**
- Geometric
- Exponential
- Normal

THE POISSON DISTRIBUTION

Recherches sur la probabilité des jugements en matière criminelle et en matière civile (1837).

The distribution was introduced to describe the theoretical number of wrongful convictions in a given country by focusing on certain random variables N that count, among other things, the number of discrete occurrences (sometimes called "events" or "arrivals") that take place during a time-interval of given length

Siméon Denis Poisson (1781–1840)
Image source:wikipedia

THE POISSON DISTRIBUTION

Soon after its introduction, the distribution was used for practical applications. For example,

In 1860, Simon Newcomb fitted the Poisson distribution to the number of stars found in a unit of space.

In 1898 Ladislaus Bortkiewicz, used the distribution to investigate the number of soldiers in the Prussian army killed accidentally by horse kicks.

In 1907, William Sealy Gosset used it to calculate the number of yeast cells used when brewing Guinness beer.

1909, A.K. Erlang, used it for the number of phone calls arriving at a call centre within a minute.

In 1946, D. Clarke used it to investigate the targeting of V-1 flying bombs on London during World War II.

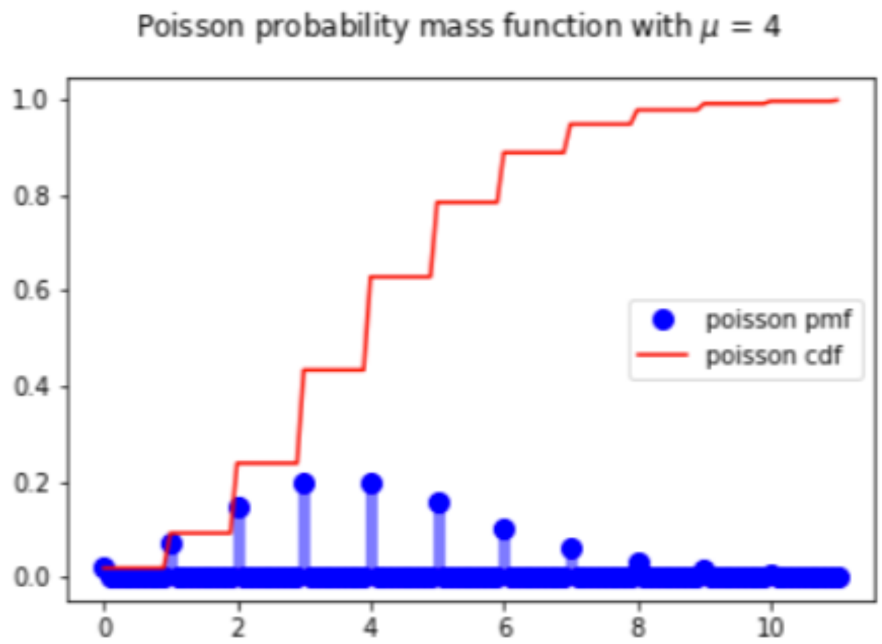
POISSON DISTRIBUTION

Definition: a random variable X having a probability mass function of the form:

$$P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

and μ can take on any positive value, is said to have a Poisson distribution.

What does it look like?



POISSON DISTRIBUTION

It is a discrete probability distribution function that describes the number of times an event that occurs infrequently in a given time period will occur.

Other examples are:

- misprints on a page in a book
- number of wrong telephone numbers dialled in a day
- number of α -particles discharged in a fixed period of time from a radioactive nuclei.
- the counts of prime numbers in short intervals (Gallagher, 1976).

CHECK IT IS A VALID PROBABILITY MASS FUNCTION

We have that

$$\sum_{\text{all } x} p(x) = \sum_{x=0}^{\infty} e^{-\mu} \frac{\mu^x}{x!}$$

$$= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

$$= e^{-\mu} e^{\mu}$$

$$= 1$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

as required.

We use the series definition of the exponential function in the next to last step.

We've not dealt with discrete variables with infinite numbers of outcomes
as far as we are concerned, they are OK as long as all properties and probabilities that we need to calculate converge as x → ∞.

EXPECTATION AND VARIANCE OF A POISSON DISTRIBUTED RANDOM VARIABLE

If X has $P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$

then

- E[X]=μ
- $\sigma^2 = \mu$

PROOF OF EXPECTATION OF POISSON DISTRIBUTED RANDOM VARIABLE

$$
\begin{aligned}
\mathrm{E}[X] &= \sum_{x=0}^{\infty} x e^{-\mu} \frac{\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x e^{-\mu} \frac{\mu^x}{x!} \\
&= e^{-\mu} \sum_{x=1}^{\infty} \frac{\mu^x}{(x-1)!} \\
&= \mu e^{-\mu} \sum_{x=1}^{\infty} \frac{\mu^{(x-1)}}{(x-1)!}
\end{aligned}
$$

then we substitute y= x – 1

$$
\begin{aligned}
&= \mu e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{(y)!} \\
&= \mu e^{-\mu} e^{\mu} \\
&= \mu
\end{aligned}
$$

PROOF OF VARIANCE OF POISSON DISTRIBUTED RANDOM VARIABLE

$$\begin{aligned}
\mathrm{E}(X^2) &= \sum_{x=0}^{\infty} x^2 e^{-\mu} \frac{\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} (x(x-1) + x) e^{-\mu} \frac{\mu^x}{x!} \\
&= \sum_{x=2}^{\infty} x(x-1) e^{-\mu} \frac{\mu^x}{x!} + \sum_{x=1}^{\infty} x e^{-\mu} \frac{\mu^x}{x!} \\
&= \mu^2 e^{-\mu} \sum_{x=2}^{\infty} \frac{\mu^{(x-2)}}{(x-2)!} + \mathrm{E}[X] \\
&= \mu^2 + \mu
\end{aligned}$$

where we substitute in y= x – 2, similarly to the expectation value.

PROOF OF VARIANCE OF POISSON DISTRIBUTED RANDOM VARIABLE

Then identify the sum as an exponential function.

We then get that

$$\begin{aligned}\sigma_X^2 &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\ &= \mu^2 + \mu - \mu^2 \\ &= \mu\end{aligned}$$

WHEN IS IT USED?

If an event with a low probability of occurrence at any instant is

1. known *to occur randomly, and*

2. *to have a mean number of occurrences μ in a given time (or space) interval*

has a Poisson distribution (approximately) with parameter μ.

An example

Government statistic for a certain country show that the average number of major industrial accidents per year for large firms is 1.1 per 5000 employees. Find the probability that there will be at least one major accident per year for firm with:

- 5000 employees
- 10000 employees

SOLUTION

Here we have a unit scale of 5000 employees and we are given the average accident rate per year is 1.1 per 5000 employees.

This should satisfy our criteria to be Poisson distributed - the accidents are rare and randomness should be a reasonable assumption.

Then we have a Poisson distribution with $\mu = 1.1$

$$P(X \geq 1) = 1 - p(0) = 1 - e^{-1.1}\frac{1.1^0}{0!} = 0.6671$$

in the interval of interest (1 year).

SOLUTION

Now we have a new scale of 10000 employees. If the accidents are rare, independent and random we'd then expect 2.2 accidents per year in a firm with 10000 employees. So

$$P(X \geq 1) = 1 - p(0) = 1 - e^{-2.2}\frac{2.2^0}{0!} = 0.8892$$

APPROXIMATION TO BINOMIAL DISTRIBUTION

One way to see how it can occur is to see how it can approximate a binomial distribution if n is large and p small.

Recall that

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

for a binomially distributed random variable.

APPROXIMATION TO BINOMIAL DISTRIBUTION

So if X is a binomially distributed random variable with parameters (n,p). If we let μ=np

$$
\begin{aligned}
P(X=x) &= \binom{n}{x} p^x (1-p)^{n-x} \\
&= \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \\
&= \frac{n!}{(n-x)!x!} \frac{\mu^x}{n^x} \left(1-\frac{\mu}{n}\right)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{n^x} \frac{\mu^x}{x!} \frac{\left(1-\frac{\mu}{n}\right)^n}{\left(1-\frac{\mu}{n}\right)^x}
\end{aligned}
$$

APPROXIMATION TO BINOMIAL DISTRIBUTION

Now for large *n* and small *p* we can make the approximations

$$\left(1-\frac{\mu}{n}\right)^n \approx e^{-\mu},$$

$$\frac{n(n-1)\cdots(n-x+1)}{n^x} \approx 1,$$

$$\left(1-\frac{\mu}{n}\right)^x \approx 1$$

Then we get that

$$P(X=x) \approx e^{-\mu}\frac{\mu^x}{x!}$$

SUMMARY OF DISTRIBUTION FUNCTIONS

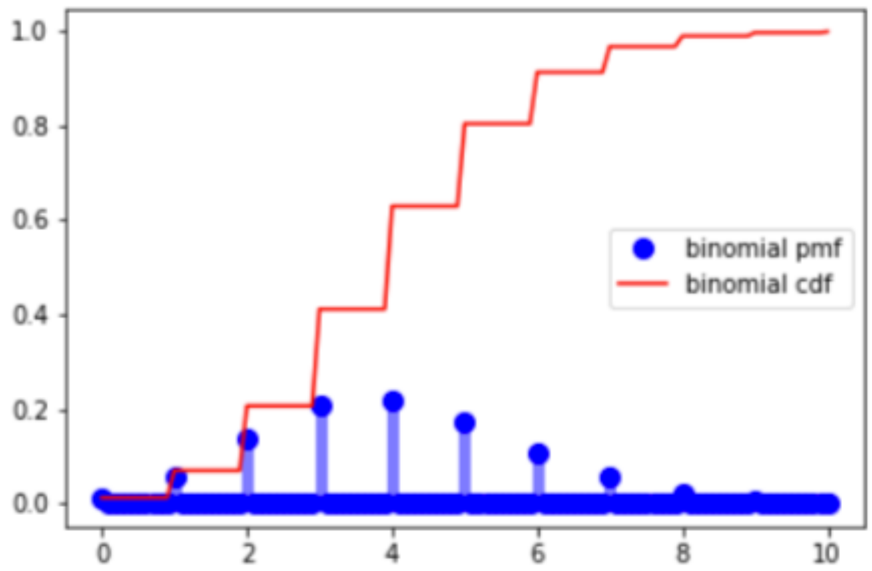
If X has

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

Then it is a **binomial distribution** with

EXPECTATION

$$\mathrm{E}[X] = np$$



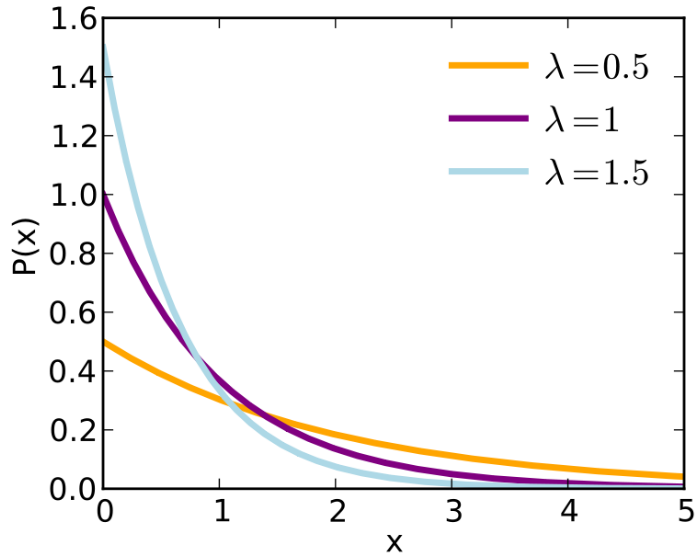
VARIANCE

$$\sigma_X^2 = E[X^2] - E[X]^2 = np(1-p)$$

SUMMARY OF DISTRIBUTION FUNCTIONS

If X has

$$P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$



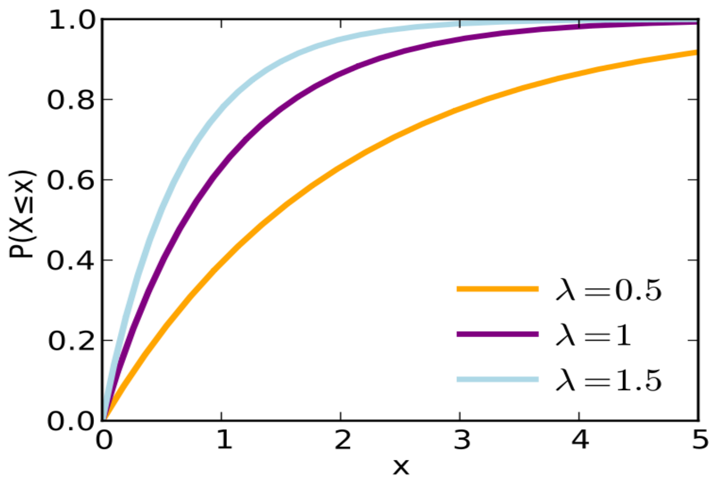
then it is a **Poisson distribution** with

EXPECTATION

$$E[X] = \mu$$

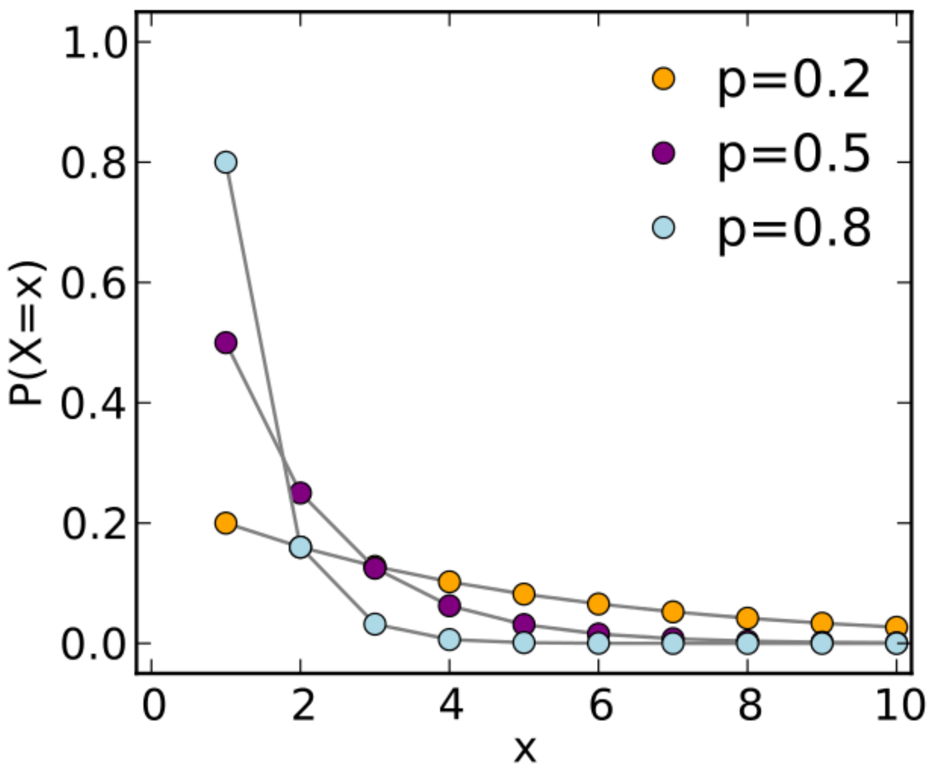
VARIANCE

$$\sigma^2 = \mu$$



Cumulative distribution function

GEOMETRIC DISTRIBUTION



GEOMETRIC DISTRIBUTION

A **discrete random variable** X with probability mass function p(x) of the form:

$$P(X = x) = p(x) = (1 - p)^{x-1}p$$

with $0 \leq p \leq 1$ is said to have a **geometric distribution**.

You can prove that it is a valid probability mass function by noting that the sum over all x is a geometric progression that can be summed up immediately.

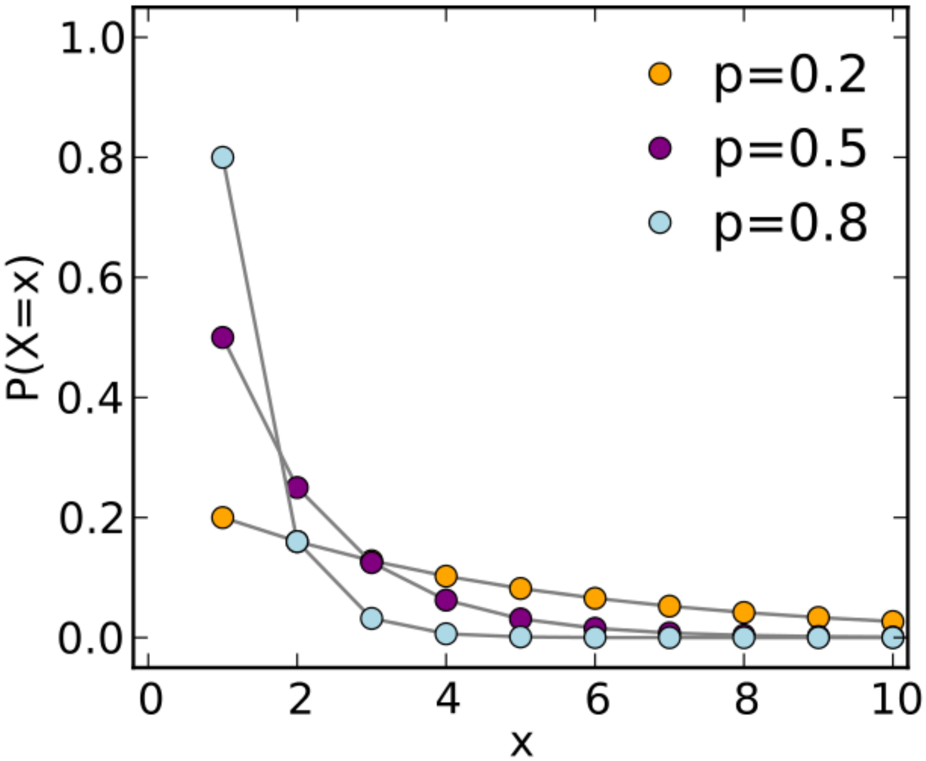
GEOMETRIC DISTRIBUTION

$$P(X = x) = p(x) = (1 - p)^{x-1} p$$

For q=(1-p) and k=x-1, we can write that

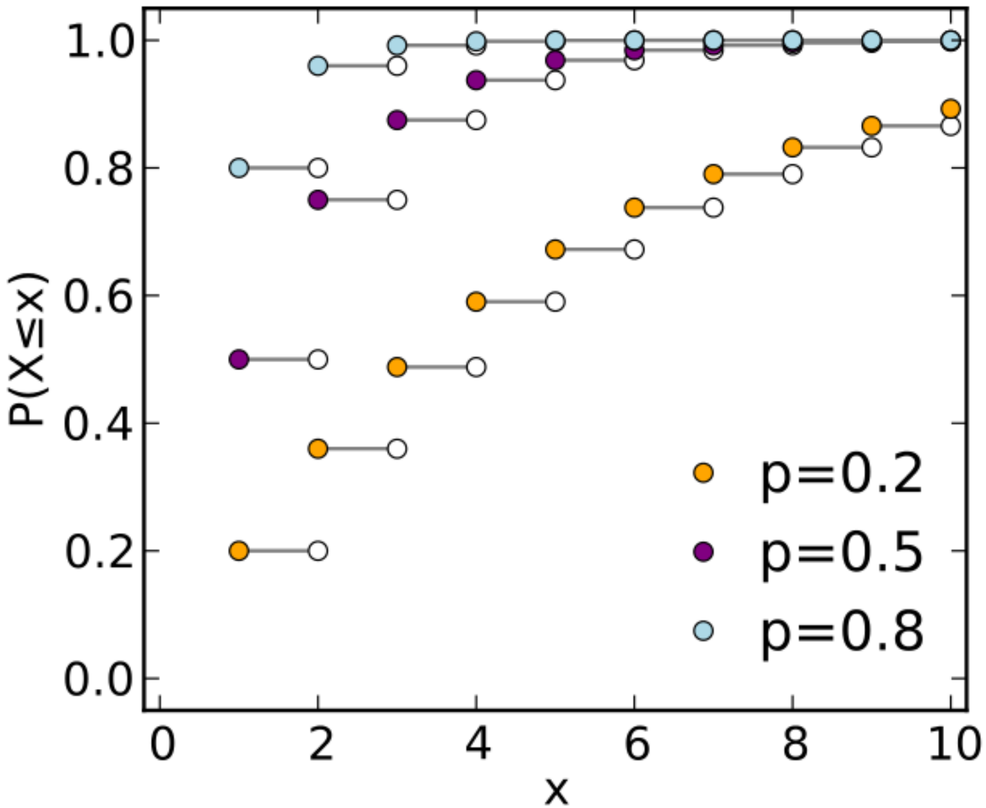
$$\sum_{x=1}^{\infty} p(1 - p)^{x-1} = \sum_{x=1}^{\infty} pq^k = \frac{p}{1 - q} = \frac{p}{1 - 1 + p} = 1$$

GEOMETRIC DISTRIBUTION



Probability distribution function

Cumulative distribution function



Source: wikipedia

GEOMETRIC DISTRIBUTION

We just give the expectation and variance here for completeness.

$$E[X] = \frac{1-p}{p}$$
$$var(X) = \frac{1-p}{p^2}$$

Note that the way in which this distribution is used can give different results:

- Some define the variable as the number of experiments until we obtain a success. In this case, we get $E[X] = \frac{1}{p}$.

- Others talk about the expected number of failures before a success. In this case, we get $E[x] = \frac{1-p}{p}$.

EXAMPLE

What is the expected number of rolls of a fair die to obtain a 1?

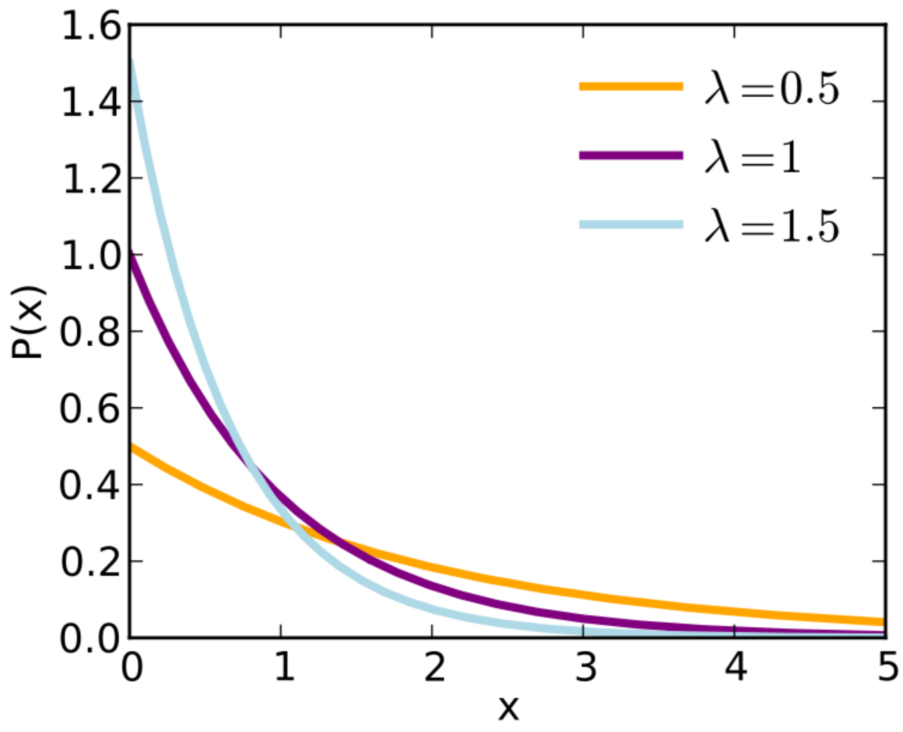
Here we formulate the question as the total number of experiments until we get a success. So, we have

$$E[X] = \frac{1}{\frac{1}{6}} = 6$$

If it had been phrased as the number of rolls before we get the 1 we'd have

$$E[X] = \frac{1 - \frac{1}{6}}{\frac{1}{6}} = 5$$

EXPONENTIAL DISTRIBUTION



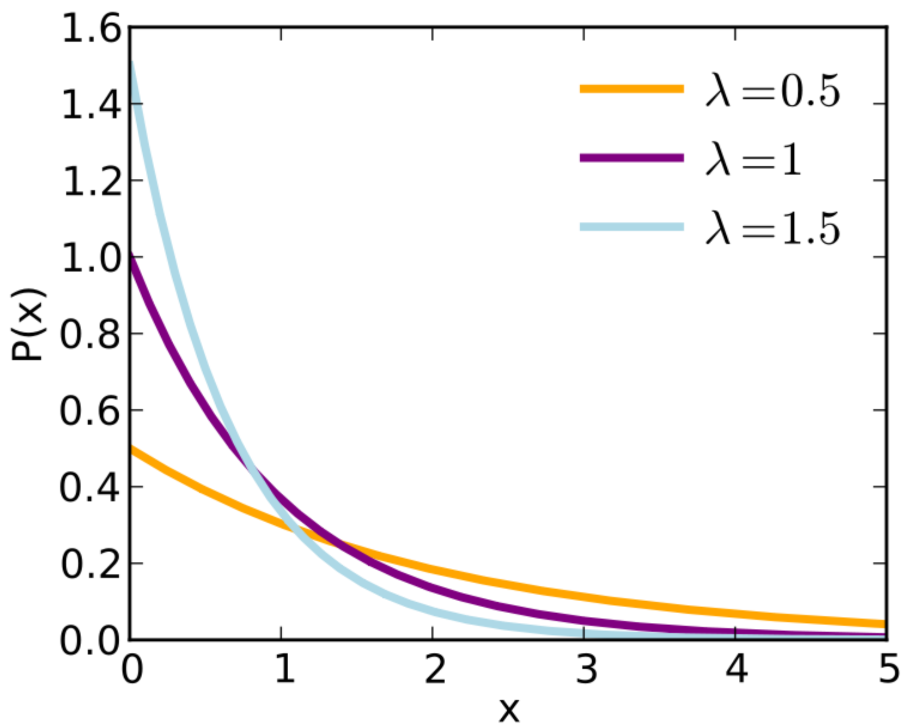
EXPONENTIAL DISTRIBUTION

The exponential distribution function of a continuous random variable y is the following

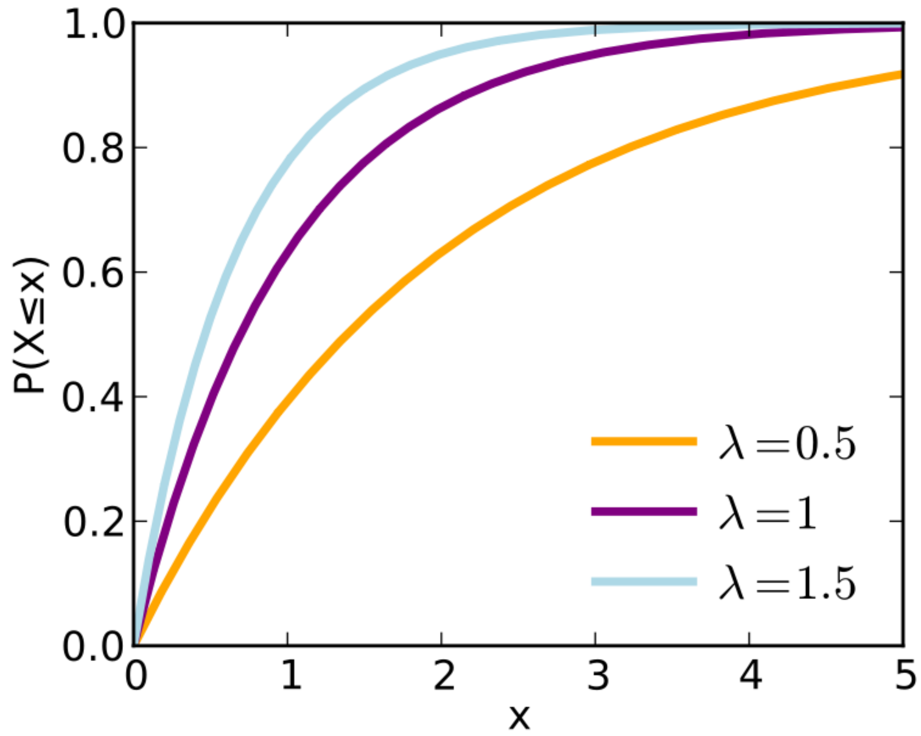
$$f_Y(y) = \beta e^{-\beta y}, y > 0$$
$$= 0 \text{ otherwise:}$$

Lets find the expectation and variance using the moment generating function of Y .

EXPONENTIAL DISTRIBUTION



Probability distribution function



Cumulative distribution function

Source: wikipedia

MOMENT GENERATING FUNCTION

There are some short cuts to calculating the moments of a distribution

The moment generating function of a random variable X is

$$m_X(t) = E[e^{tX}], \qquad -\infty < t < \infty$$

This function can be used to generate (hence the name) the ***moments of a distribution.***

Recalling that

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots$$

the Taylor series of e^{tx} about x=0

MOMENT GENERATING FUNCTION

Then

$$
\begin{aligned}
m_X(t) &= \mathrm{E}[e^{tX}] \\
&= \mathrm{E}\left[1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right]
\end{aligned}
$$

and again using the linear properties of E[] we get

$$
\begin{aligned}
m_X(t) &= 1 + \mathrm{E}[tX] + \frac{t^2}{2!}\mathrm{E}[X^2] + \frac{t^3}{3!}\mathrm{E}[X^3] + \cdots \\
&= 1 + tm_1 + \frac{t^2}{2!}m_2 + \frac{t^3}{3!}m_3 + \cdots
\end{aligned}
$$

MOMENT GENERATING FUNCTION

If we now take derivatives with respect to *t* we get

$$\frac{\mathrm{d}m_X(t)}{\mathrm{d}t} = m_1 + tm_2 + \frac{t^2}{2!}m_3 + \cdots$$
$$\frac{\mathrm{d}^2 m_X(t)}{\mathrm{d}t^2} = m_2 + tm_3 + \frac{t^2}{2!}m_4 + \cdots$$

Finally setting t=0 all terms in the series will be equal to zero except the first

$$\left.\frac{\mathrm{d}m_X(t)}{\mathrm{d}t}\right|_{t=0} = m_1$$
$$\left.\frac{\mathrm{d}^2 m_X(t)}{\mathrm{d}t^2}\right|_{t=0} = m_2$$

Etc.

EXPONENTIAL DISTRIBUTION

The moment generating function of Y is

$$
\begin{aligned}
m_Y(t) &= \mathrm{E}[e^{tY}] \\
&= \int_0^\infty e^{ty}\beta e^{-\beta y}\,\mathrm{d}y \\
&= \int_0^\infty \beta e^{-y(\beta - t)}\,\mathrm{d}y \\
&= \frac{\beta}{\beta - t}
\end{aligned}
$$

$$m^{(1)}(0) = \frac{\mathrm{d}(\mathrm{m_t}(0))}{\mathrm{dt}} = \frac{\beta}{(\beta - t)^2} = \beta^{-1}$$

to find the moments we take the derivative with respect to t then set *t=0*

$$
\begin{aligned}
\mu_Y &= m^{(1)}(0) = \beta^{-1} \\
m_2 &= m^{(2)}(0) = 2\beta^{-2}
\end{aligned}
$$

so

$$\sigma_Y^2 = m_2 - \mu^2 = \beta^{-2}$$

EXPONENTIAL DISTRIBUTION

The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process.

For example:

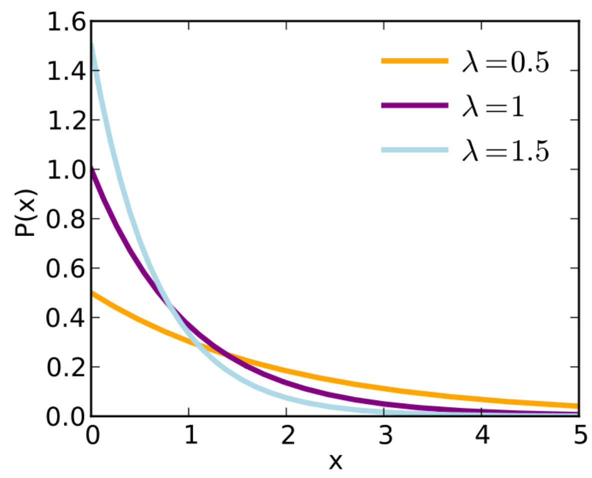
- The time until a radioactive particle decays, or the time between clicks of a Geiger counter
- The time it takes before your next telephone call
- How long it takes for a bank teller etc. to serve a customer.
- Distance between DNA mutation in a gene.

SUMMARY

If X has

$$f_Y(y) = \beta e^{-\beta y}, y > 0$$
$$= 0 \text{ otherwise:}$$

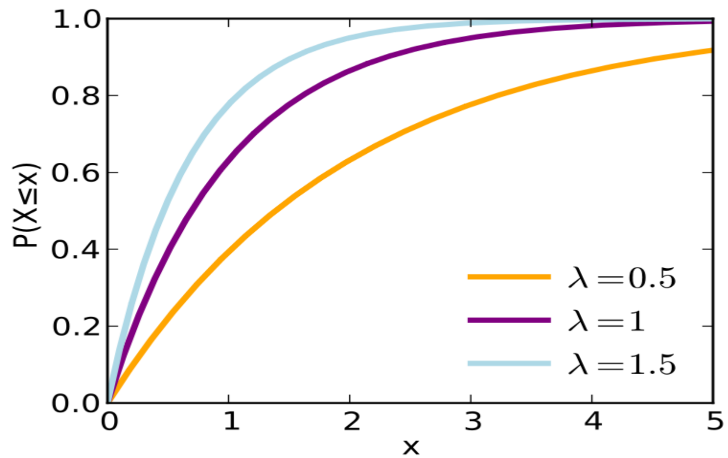
Then it is a **exponential distribution** with



PMF

EXPECTATION

$$\mu_Y = m^{(1)}(0) = \beta^{-1}$$



Cumulative distribution function

VARIANCE

$$\sigma_Y^2 = m_2 - \mu^2 = \beta^{-2}$$

NORMAL or **GAUSSIAN** or **GAUSS** or **LAPLACE–GAUSS DISTRIBUTION**

Carl Friedrich Gauß (1777–1855), painted by Christian Albrecht Jensen. (Wikipedia)

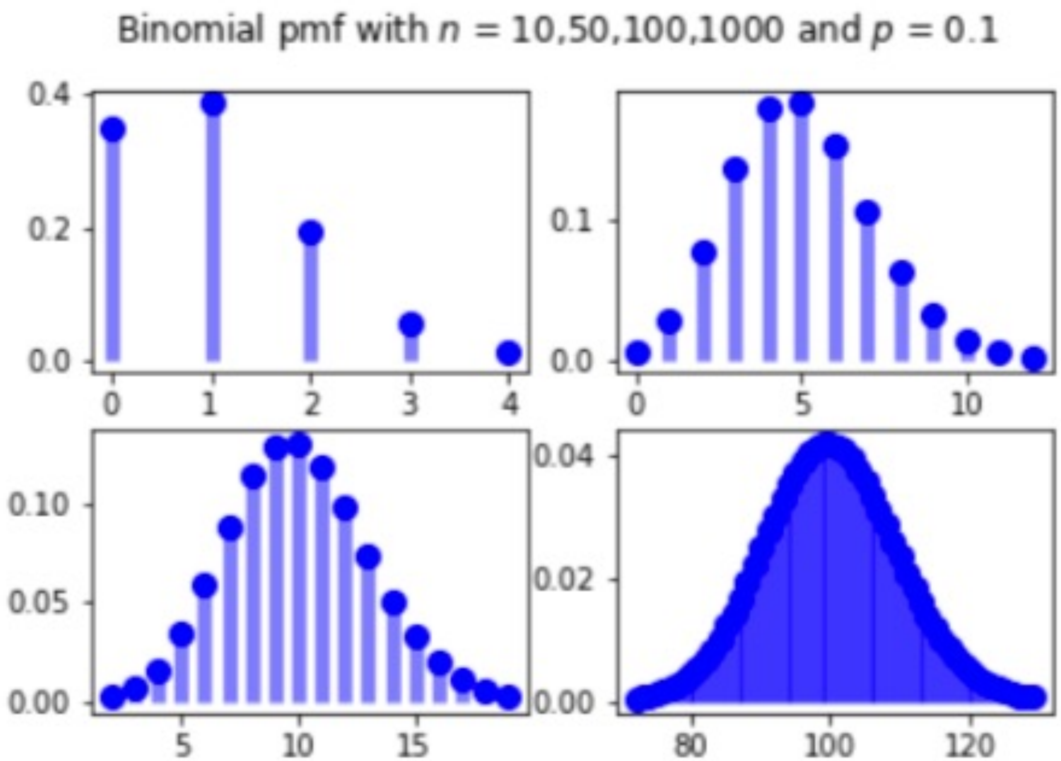
DISTRIBUTION OF LARGE DATA SET

lets have a look at these distributions as N gets large

DISTRIBUTION OF LARGE DATA SET

First the Binomial distribution

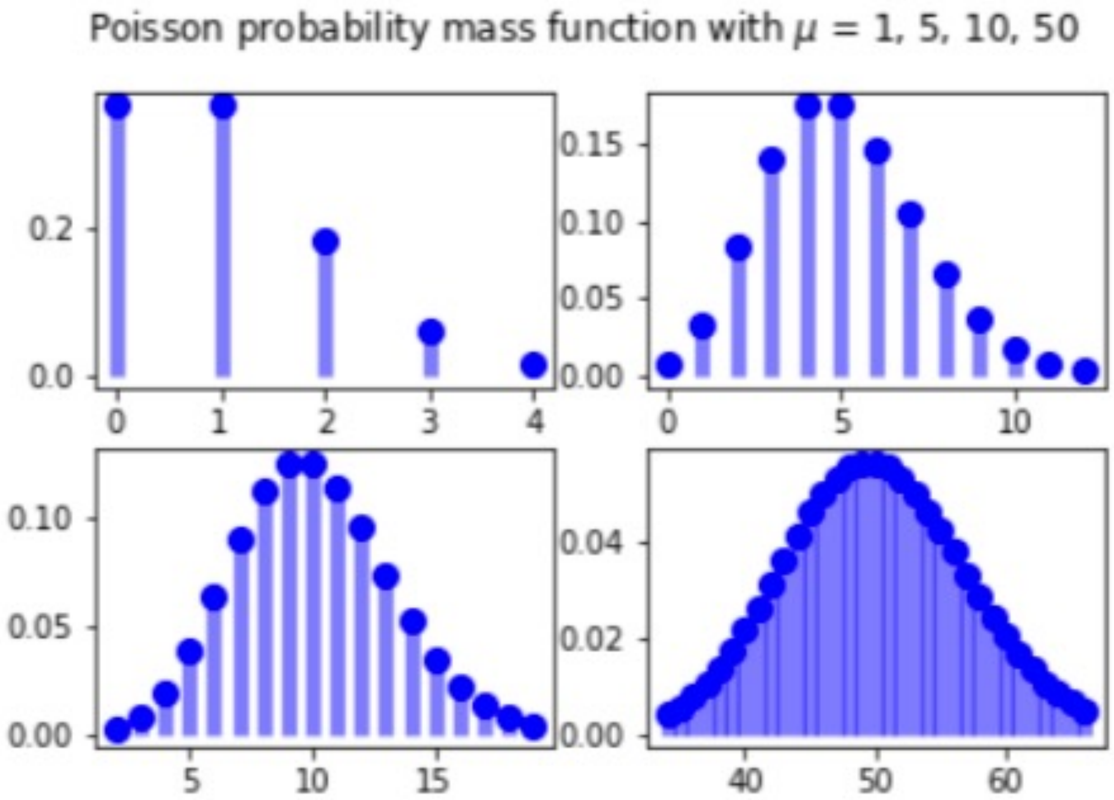
$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, (x = 0, 1, \ldots, n-1, n)$$



DISTRIBUTION OF LARGE DATA SET

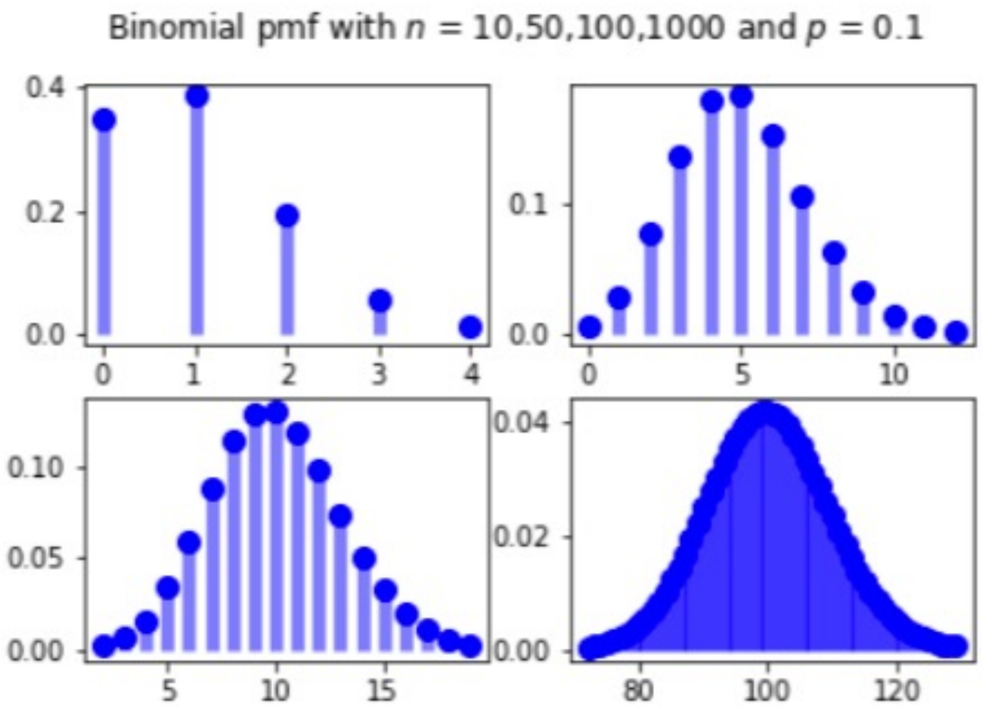
And now for the Poisson distribution

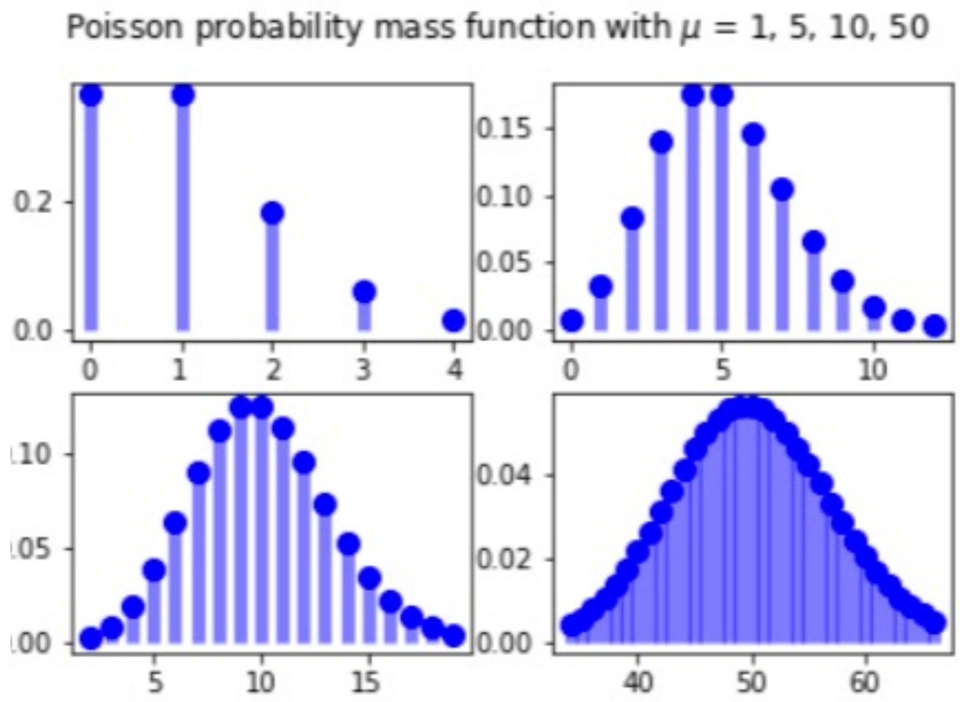
$$p(x) = e^{-\mu}\frac{\mu^x}{x!},$$



DISTRIBUTION OF LARGE DATA SET

We see that both tend towards a bell shaped curve as n increases





(you can also see that the binomial and Poisson look quite similar, Poisson should be a good approx. when $\mu = np$)

NORMAL DISTRIBUTION

This brings us to the most important distribution in all of probability and statistics, the normal distribution.

The name gives it away (though this probably not the reason for the name) – this is the distribution that you normally expect to see:

- Peoples' heights or weights,
- velocities of atoms ...
- if there are enough of something, it will be normally distributed!

This property, that most distributions tend to the normal distribution as the sample/population gets large, has the big name:

Central Limit Theorem

We'll look in detail at the normal distribution and some distributions derived from it   now.

Definition

We say that X is a normal random variable, or that X is normally distributed, with parameters μ and σ^2 if the probability density function of X is given by

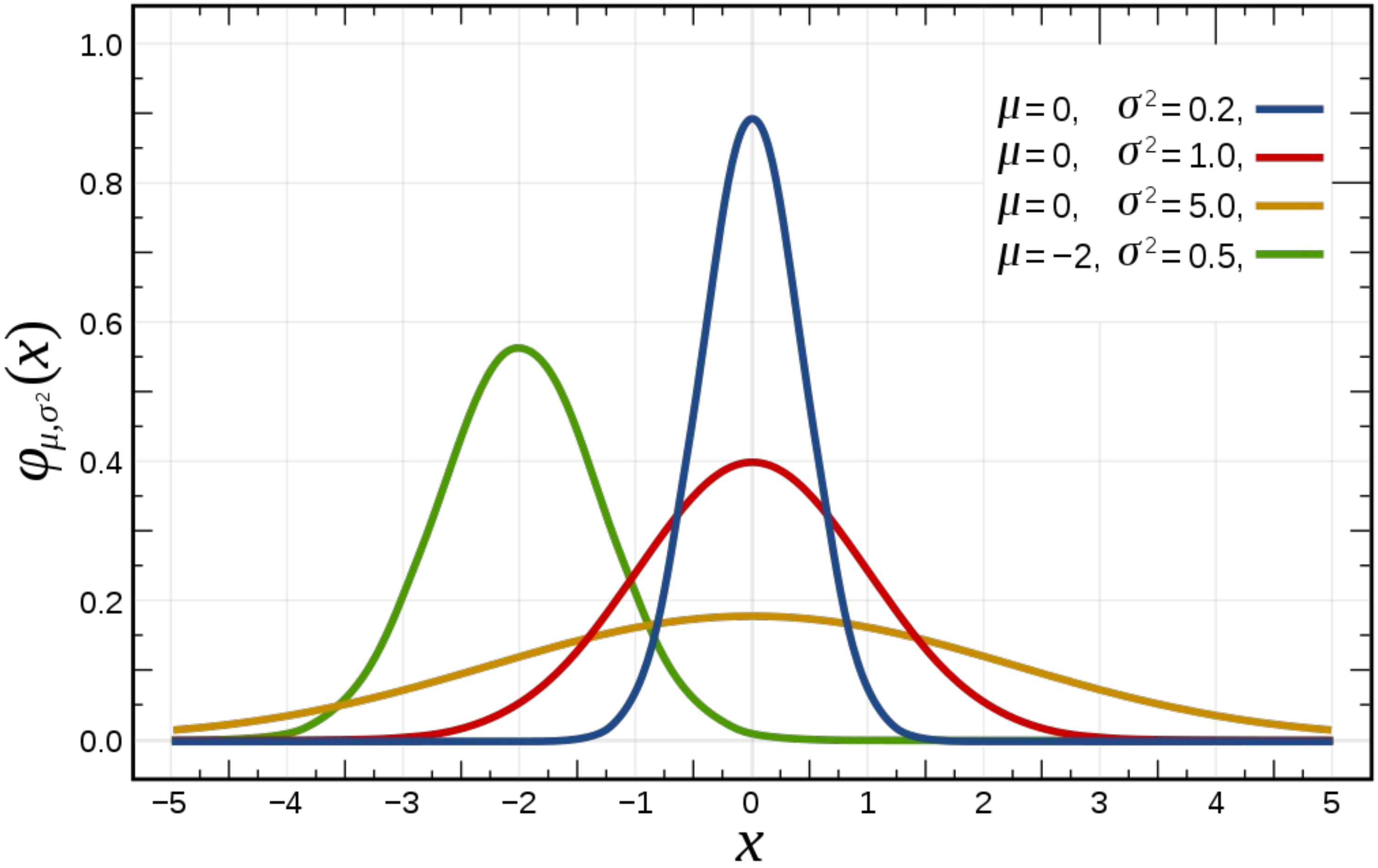
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This function is also called a **Gaussian function.**

We'll see that the parameters **μ** and **σ** control the location of the centre of the distribution and its width. μ defines the centre of the distribution, and as the distribution is symmetric about the centre, its mean and median.

We'll also see that σ is the standard deviation of this pdf.

Definition



Source: wikipedia

CHECK IF IT IS A VALID PDF

It is not at all obvious that

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

is a valid probability density function.

The integration

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{dx}$$

requires a few tricks. You only need the result, but it is good that you see a sketch of how the integral can be done.

CHECK IF IT IS A VALID PDF

First, we define a new variable $y = \frac{x-\mu}{\sigma}$, we then see that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{dx} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{\mathrm{y}^2}{2}} \mathrm{dy}$$

So we want to show that

$$I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \mathrm{dy} = \sqrt{2\pi}$$

CHECK IT IS A VALID PDF

Now lets instead consider the square of I

$$I^2 = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,\mathrm{dy} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\,\mathrm{dx}$$

note the two variables are different otherwise it would not be the simple product of the two integrals, which after all are only simple numbers.

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}\,\mathrm{dy}\,\mathrm{dx}$$

The point of squaring the integral is that we can now change to polar coordinates (remember calculus last term), though this goes beyond what you have yet done with multiple integrals.

CHECK IT IS A VALID PDF

We let $x = rcos\theta$ and $y = rsin\theta$. We also have that $dxdy = rd\theta dr$. Then we find that

$$
\begin{aligned}
I^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}\,\mathrm{dy}\,\mathrm{dx} \\
&= \int_{0}^{\infty}\int_{0}^{2\pi} e^{-\frac{r^2}{2}}\, r\,\mathrm{d}\theta\,\mathrm{dr} \\
&= 2\pi\int_{0}^{\infty} re^{-\frac{r^2}{2}}\,\mathrm{dr}
\end{aligned}
$$

Surprisingly this integral is much easier than I .

CHECK IT IS A VALID PDF

$$I^2$$

$$= 2\pi \int_0^{\infty} re^{--\frac{r^2}{2}}\,\mathrm{dr}$$

We can substitute $s = -r^2$ to carry out the definite integral, and $ds = -2rdr$, leading to

$$I^2 = \pi \int_0^{\infty} -e^{-\frac{s}{2}}\,\mathrm{ds}$$

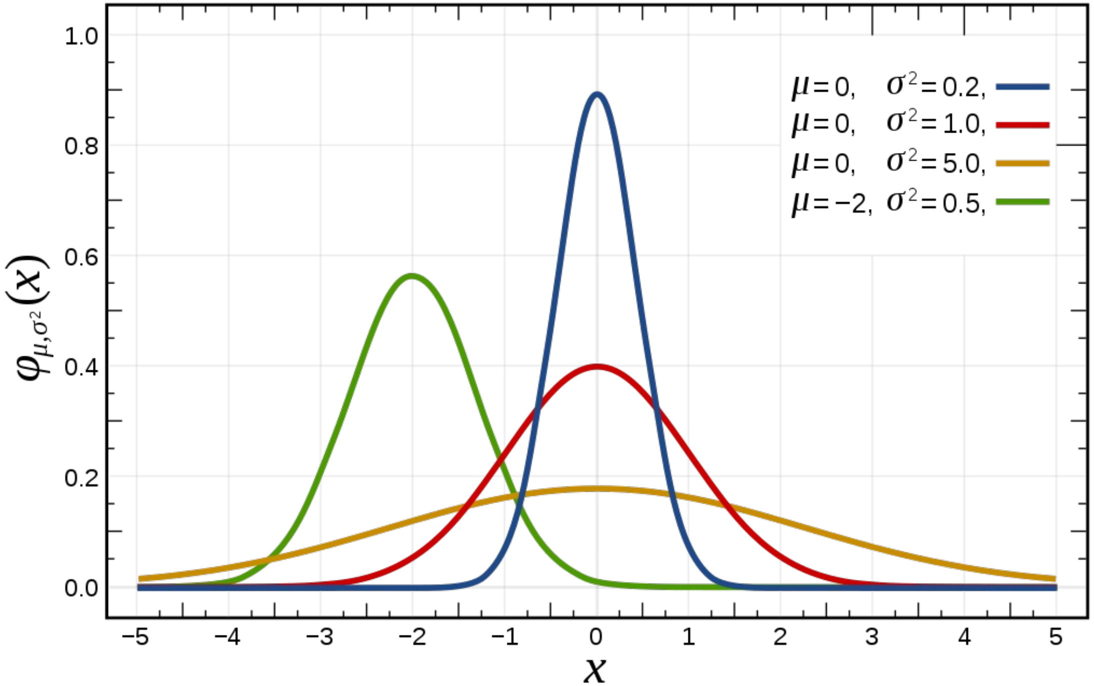
$$I^2 = 2\pi e^{\frac{s}{2}} \Big|_{-\infty}^{0}$$

$$= 2\pi$$

Hence $I = \sqrt{2\pi}$ and we have proved that $f(x)$ is a valid density functior

EXPECTATION AND VARIANCE OF THE STANDARD NORMAL DISTRIBUTION

Find the expectation and variance of a standard normally distributed random variable Z with parameters μ = 0 and σ = 1.



The expectation of Z is given by

$$\begin{aligned} \mathrm{E}[Z] &= \int_{-\infty}^{\infty} z f_Z(z) \mathrm{dz} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}} \mathrm{dz} \\ &= 0 \end{aligned}$$

EXPECTATION AND VARIANCE OF THE STANDARD NORMAL DISTRIBUTION

The variance is given by (note $E[Z] = 0$)

$$\begin{aligned} \sigma_Z^2 &= \mathrm{E}[Z^2] - (\mathrm{E}[Z])^2 \\ &= \mathrm{E}[Z^2] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 \mathrm{e}^{-\frac{z^2}{2}} \mathrm{d}z \end{aligned}$$

We can integrate by parts to obtain with $u = z; dv = ze^{-\frac{z^2}{2}}$

EXPECTATION AND VARIANCE OF THE STANDARD NORMAL DISTRIBUTION

$$
\begin{aligned}
\sigma_Z^2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \left(- z e^{-\frac{z^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \mathrm{d}z \right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \mathrm{d}z \\
&= 1
\end{aligned}
$$

using the property of the normal pdf.

STANDARD NORMAL VARIABLE

The standard normal variable is a normally distributed random variable with $\mu = 0$ and $\sigma = 1$.

The reason this is important is that through a change of variables we can map any other normally distributed variable onto the standard case.

Proposition: if X is normally distributed with parameters μ and σ then

$$X = \mu + \sigma Z$$

where **Z** is a variable with the standard normal distribution.

STANDARD NORMAL VARIABLE

Now we can use our previous results on the properties of expectations and variances to show that the expectation and variance of the general normally distributed random variable $X = \mu + \sigma Z$ is given by

$$\begin{aligned} \mathrm{E}[X] &= \mathrm{E}[\mu + \sigma Z] \\ &= \mathrm{E}[\mu] + \mathrm{E}[\sigma Z] \\ &= \mu + \sigma \mathrm{E}[Z] \\ &= \mu \end{aligned}$$

$$\begin{aligned} \mathrm{var}(X) &= \mathrm{var}(\mu + \sigma Z) \\ &= \mathrm{var}(\sigma Z) \\ &= \sigma^2 \mathrm{var}(Z) \\ &= \sigma^2 \end{aligned}$$

The important point is we basically never have to work with a general normally distributed random variable, we just transform to Z and do everything in terms of Z and its distribution.