The Binomial Distribution

- The binomial distribution is a fundamental result from probability theory.
- This distribution function isn’t encountered often in experimental physics.
- However **it can be used to explain why measurements almost always have a Gaussian distribution of values**.
- The name derives from the similarity of the expression for the distribution to the binomial expansion formula.

For an experiment that has multiple **discrete** outcomes, the probability of finding a certain result is given by the binomial distribution.

- E.g. consider tossing a coin 4 times, the probability of obtaining heads 3 times is described by a binomial distribution.

$$Prob(\nu\ successes\ in\ n\ trials) = B_{n,p}(\nu)$$

Probability of “success” in any single trial (0-1).

No. of trials

$$= \frac{n!}{\nu!\,(n-\nu)!}\,p^{\nu}(1-p)^{n-\nu}$$

Target value

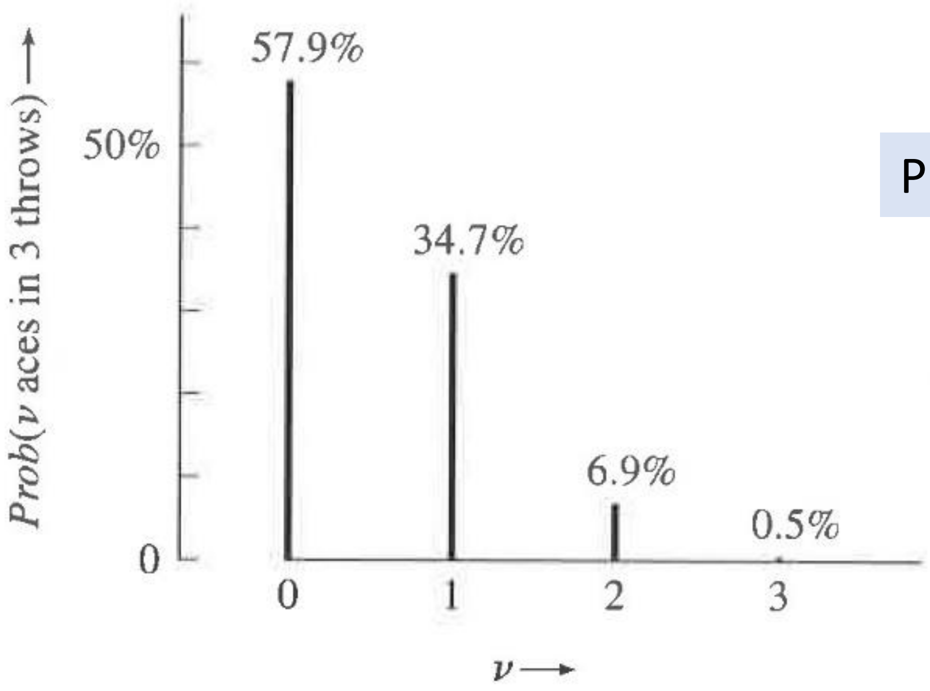
$$= \binom{n}{\nu} p^{\nu}(1-p)^{n-\nu}$$

Often written in this form (Note 0! =1)

E.g. Consider throwing a six-sided die 3 times and asking what are the probabilities of rolling a six 0, 1, 2, or 3 times?

- Here $n = 3$ and $p = 1/6$.

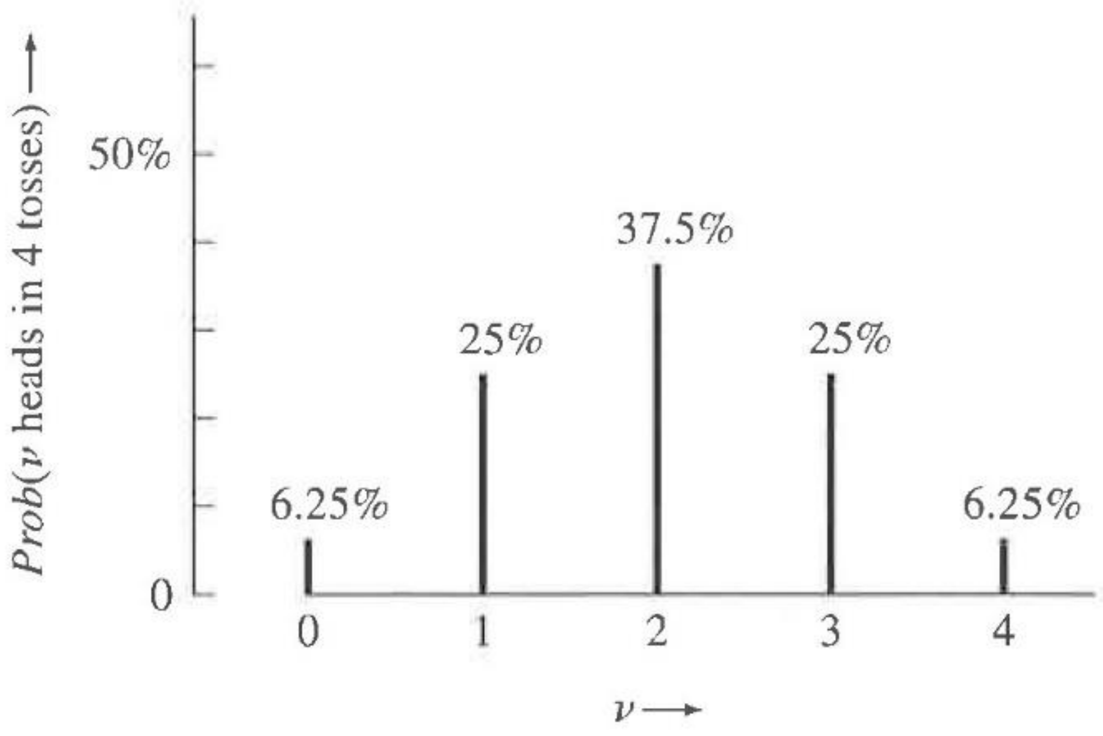
$Prob(0\ sixes\ in\ 3\ throws)$ =0.579
$Prob(1\ in\ 3)$ =0.347
$Prob(2\ in\ 3)$ =0.0694
$Prob(3\ in\ 3)$ =0.0046



Probabilities of rolling ν sixes in 3 rolls of the die.

- A feature of the binomial distribution is that, in general, the distribution isn't symmetric in shape.

- In the case of $p = 1/2$ (e.g. the chance of finding heads when tossing a coin once), the distribution **is** symmetric about the most probable outcome.



Probabilities of finding n heads when throwing 4 coins. ($n = 4, p = 1/2$)

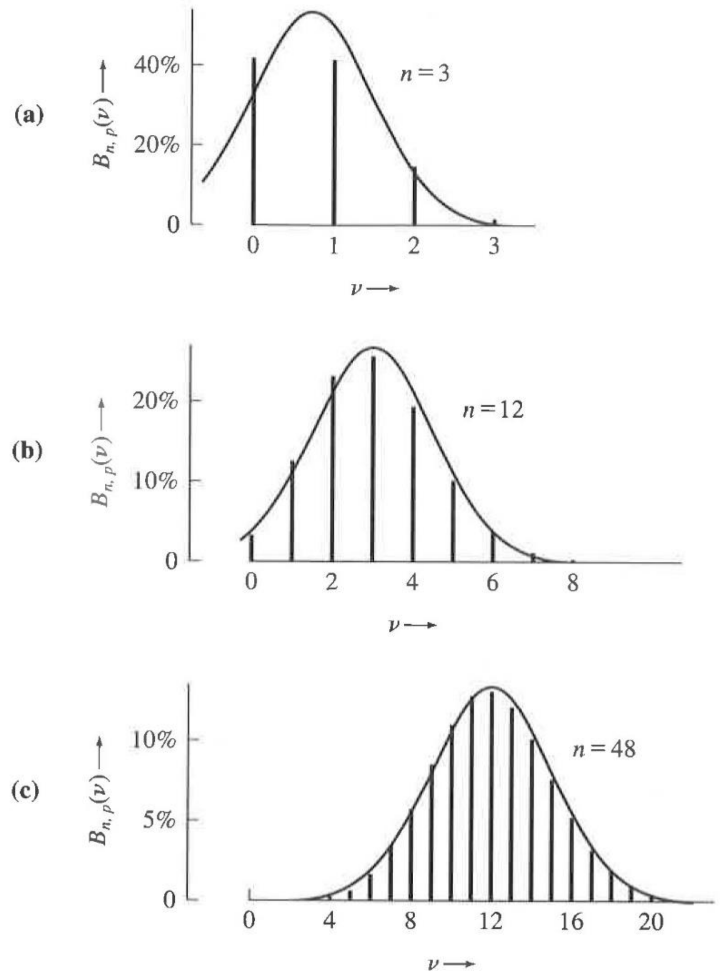
Some properties of the binomial distribution:

Average value (of the number of successes): $\bar{\nu} = np$

standard deviation: $\sigma = \sqrt{np(1-p)}$.

Gaussian approximation to the binomial distribution

When n is large, the discrete binomial distribution is closely approximated by a continuous Gaussian function (irrespective of p). (see reference in the notes).



- This is a consequence of the central limit theorem.

This can be demonstrated in a phenomenological way.

- Compare binomial distributions for $p = 1/4$ superimposed with a Gaussian distribution (same mean and σ) for increasing values of n.

The binomial distribution (discrete line plots) for $p = 1/4$ for a) n=3, b) n=12, c) n=48.
The solid curve in each plot is a Gaussian distribution.

As n becomes large the distribution becomes Gaussian (irrespective of p)

The Gauss distribution for random errors

We will see that if a measurement is subject to many small, random contributions, the results will be given by a Gaussian distribution.

- Gaussian distribution also known as the normal distribution.

Suppose we measure a quantity x that has a true value X.

Assumptions:

- Measurements aren't subject to any systematic errors (e.g. a slow stopwatch).
- There are n sources of random error (e.g. reaction time, reading parallax error, air currents, etc).
- Each of these sources of random error are of the same sized ε. (for simplicity).

- Consequently, each source of error can push our measured result up or down by ε . Assume each possibility has equal probability (i.e., p=1/2)

If we had one error source:

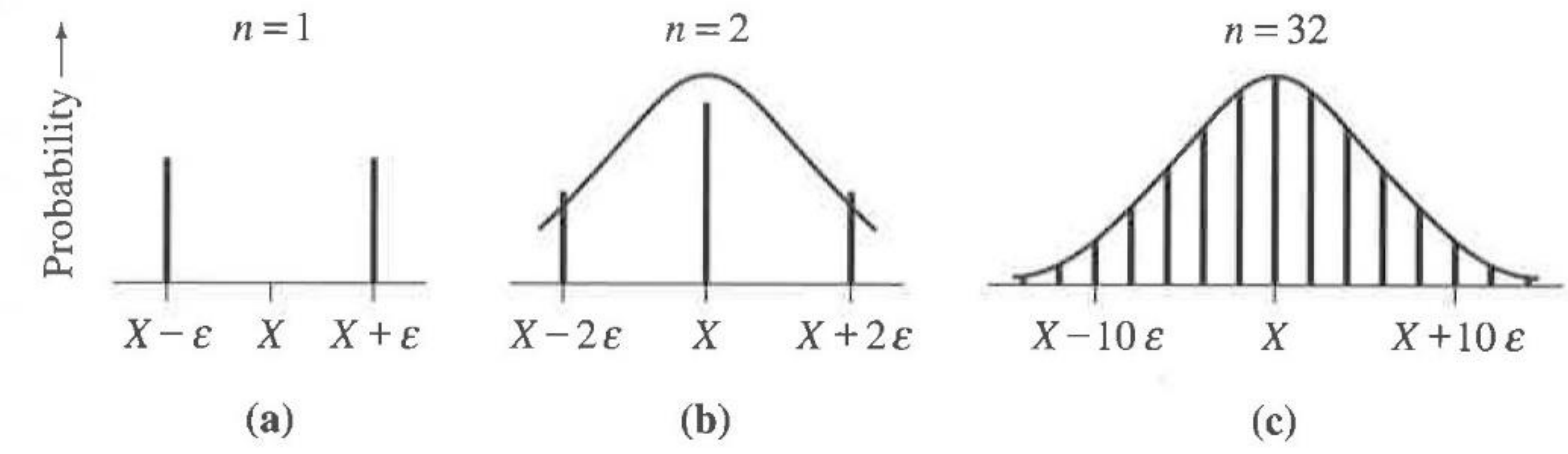
Either get: $x = X + \varepsilon$ or $x = X - \varepsilon$

For two error sources:

$$x = X - 2\varepsilon, \quad x = X + 2\varepsilon \quad \text{or} \quad x = X$$

(both negative)

(if one error was positive and one negative)



The distribution of measured values assuming n random errors of magnitude ε, for n =1, 2 and 32. The continuous curves superimposed are Gaussians with the same centre and width.

- If there are n sources of error the measured value can range between $x = X \pm n\varepsilon$.
- The probability of a given result (say ν positive errors) is given by the binomial distribution.

If there are a large number of sources of error: $\Rightarrow$ The measurements will be normally distributed (a Gaussian distribution function).

- This is a consequence of the binomial distribution being well approximated by a Gaussian for large n.

In experimental work, it is customary to assume that uncertainties in measured quantities are normally distributed (as this is commonly observed).

- This assumption is central to the familiar uncertainty propagation formulae that are used in physics.

The main exception to this rule is discussed next.

The Poisson distribution

The Poisson distribution describes the results of experiments in which individual events occur at random, **but at a definite average rate**.

A typical example is decay from a radioactive sample. ⇨

- Here, there is a well-defined average rate at which decays occur (the decay constant).
- However, the timing between individual decays is randomly distributed.
- The number of decays per minute measured in a series of identical time trials is given by a Poisson distribution.

The Poisson distribution actually is a limiting case of the binomial distribution.

To see this:

Say we do a radioactive counting experiment many times (for a fixed time), we will see (in general) a different number of counts each time.

- The probability of seeing a particular number of counts, ν, can be calculated (as expected) using the binomial distribution.
- Here n here is the number of atoms in the sample and p is the probability of an atom decaying in the time of the counting window.

Consequently:

- p is very small (for radioactive decay, p may be $\sim 10^{-20}$ per s).
- n is very large (n may be $> 10^{20}$ atoms).

Thus we can make a simplification to the binomial distribution:

$$\Rightarrow \quad P_\mu(\nu) = e^{-\mu}\frac{\mu^\nu}{\nu!}.$$

average count

The probability, of finding counts ν in the time interval over which the average count, μ is specified.

Also, we have:

$$\mu = np$$

As expected as p is the probability of one atom decaying in the expt. time.

An example:

A sample of Thorium emits alpha particles at an average rate of 1.5 per minute.
What does the distribution of counts look like assuming a 2 minute count detection window?

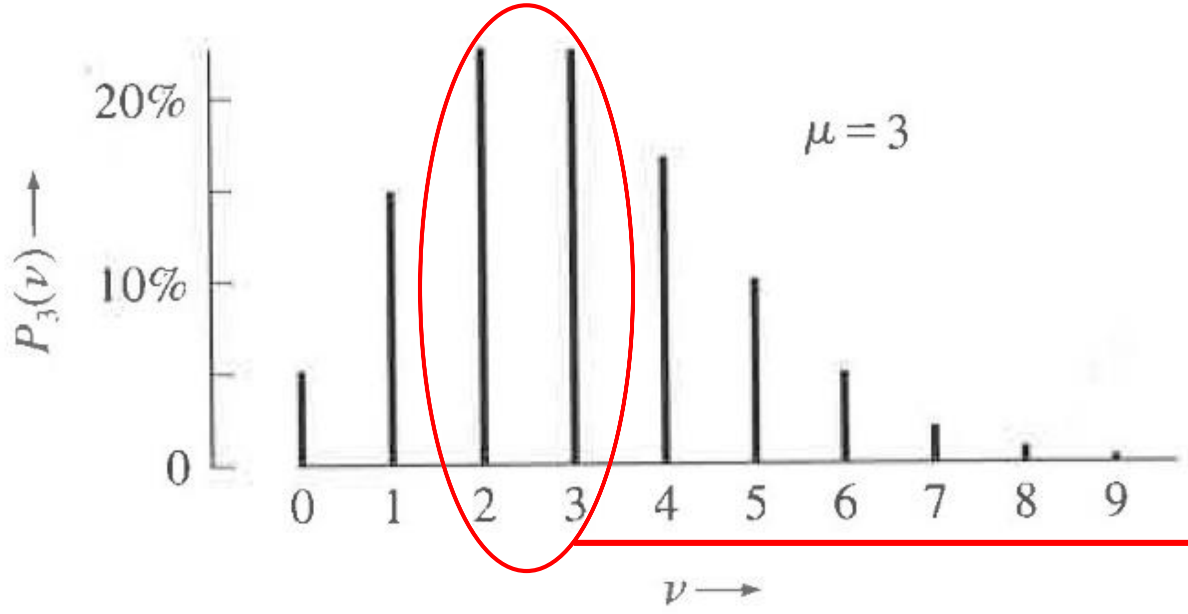
The average count value expected: $\mu = 3$ (3 counts in 2 mins)

Choose to plot the probability $P_\mu(\nu)$ for $\nu = 0, 1, 2, 3, \ldots 9$.

Calculate the probability of obtaining 3 counts, i.e., $\nu = 3$: $\Rightarrow$

$$Prob(3\ particles) = e^{-3}\frac{3^3}{3!}$$

$$= 0.224 \qquad (22.4\%)$$



The Poisson distribution for observing events in a counting experiment when the expected average count is 3.

The probability for 2 and 3 counts is actually identical!

Properties of the Poisson distribution

average count

$\boldsymbol{\mu}$

- This actually is one of the “input parameters” to define the distribution.
- The value that would be obtained if an infinite number of trials would be performed.

standard deviation

$\boldsymbol{\sigma}$

$$\sigma = \sqrt{Np}$$

From the binomial distribution properties as $(1 - p) \approx 1$.

$$\sigma = \sqrt{\mu}\,.$$

A Poisson distribution with a mean count μ has a standard deviation of $\sqrt{\mu}$.

If we do a counting experiment **once** and get the value ν what does this tell us about the actual value of μ?

- The most likely value for μ is actually the observed counts, ν.

Consequently, the uncertainty in the count value ν is just $\pm\sqrt{\nu}$

The "square-root rule" for counting experiments:

In experiments where events occur randomly, but with a definite average rate, the uncertainty in the number of counts n in a measurement is just $\sqrt{n}$.

Thus, the number of counts should be quoted as:

$$n \pm \sqrt{n}\,.$$

Typical applicability:

- The number of counts in a radioactive decay counting experiment.
- Noise of an individual pixel on a CCD (shot noise).
- The number of photons detected by a photomultiplier or avalanche photodiode.

An example:

A student monitors the decay of a thorium sample for 30 minutes and observes 49 alpha particles. What is the particle emission rate per minute?

$$\begin{aligned} n &= 49 \pm \sqrt{49} \\ &= 49 \pm 7\ particles \end{aligned}$$

Uncertainty based on Poisson statistics

To get the number per minute (a rate):

$$R = \frac{n}{30} = \frac{49 \pm 7}{30} = 1.6\ \pm 0.2\ \text{ particles/min}$$

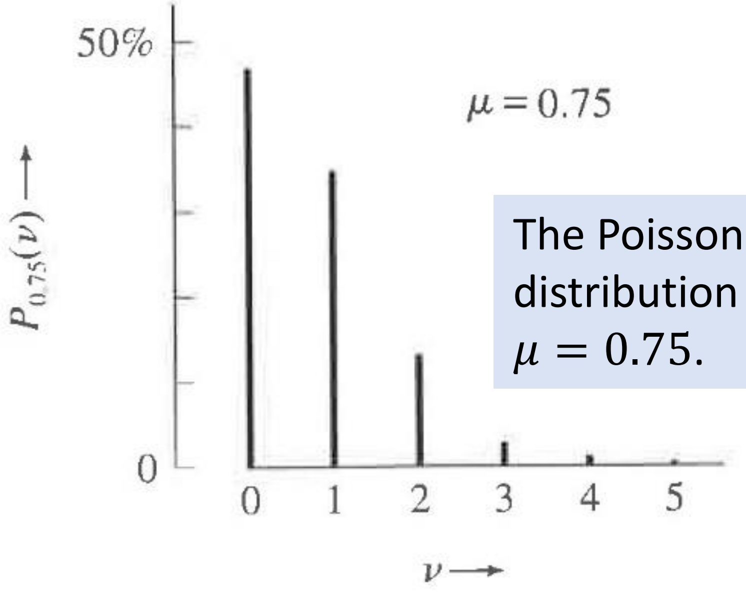
Exercise

Gaussian approximation to the Poisson distribution

Under certain conditions, we can approximate the Poisson distribution with a Gauss function.

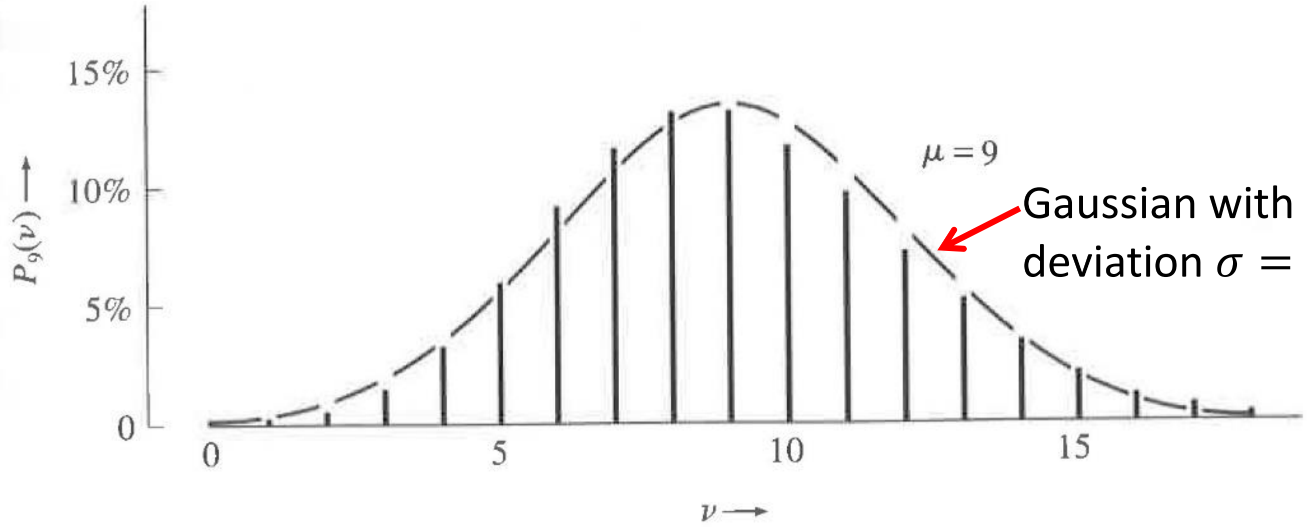
- The advantage of this is that the Gaussian is a continuous function.

- For low mean count values, the Poisson distribution is highly asymmetric.
- As the count values increase, the distribution becomes more symmetric and begins to resemble a Gaussian.



The Poisson distribution for $\mu = 0.75$.

The Poisson distribution for $\mu = 9$ (vertical lines).



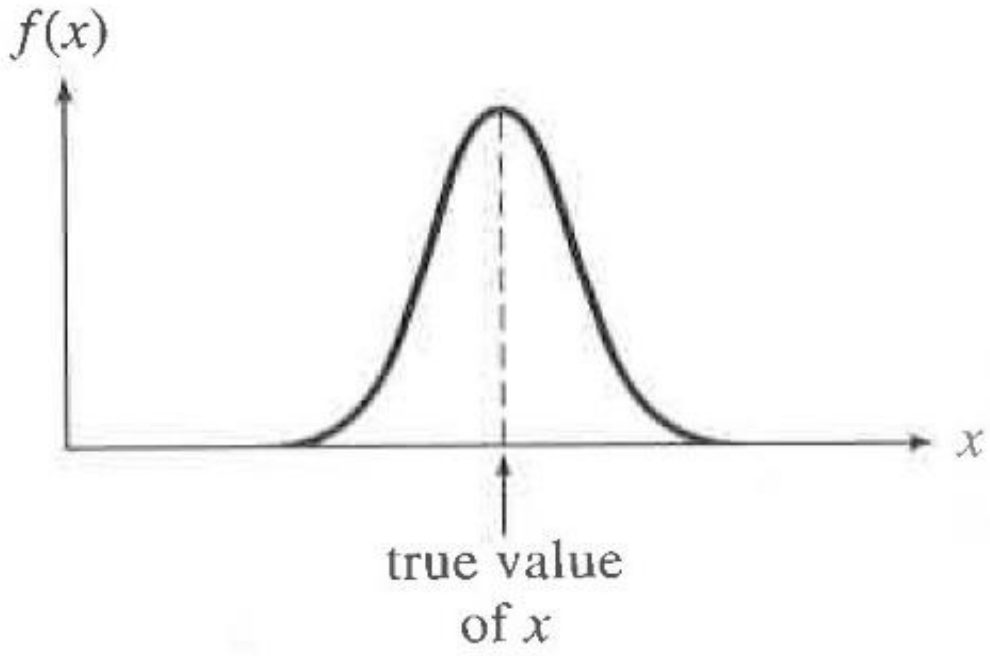
Gaussian with a standard deviation $\sigma = 3$ (i.e., $\sqrt{\mu}$).

A Gaussian distribution can be used to replace the Poisson distribution when the mean count value, μ, becomes large (say $\mu \geq 10$).

The normal (Gaussian) distribution

In many situations you can expect the distribution of values to be well approximated by a Gauss function.

- Any time multiple, random, influences apply to measurements, you can expect the distribution of the results to conform to a normal distribution.
- The most probable value corresponds to the true value of the parameter.



The limiting distribution for a measurement subject to many small random errors is a normal distribution.
The bell shaped distribution is centred on the true value of the measured quantity.

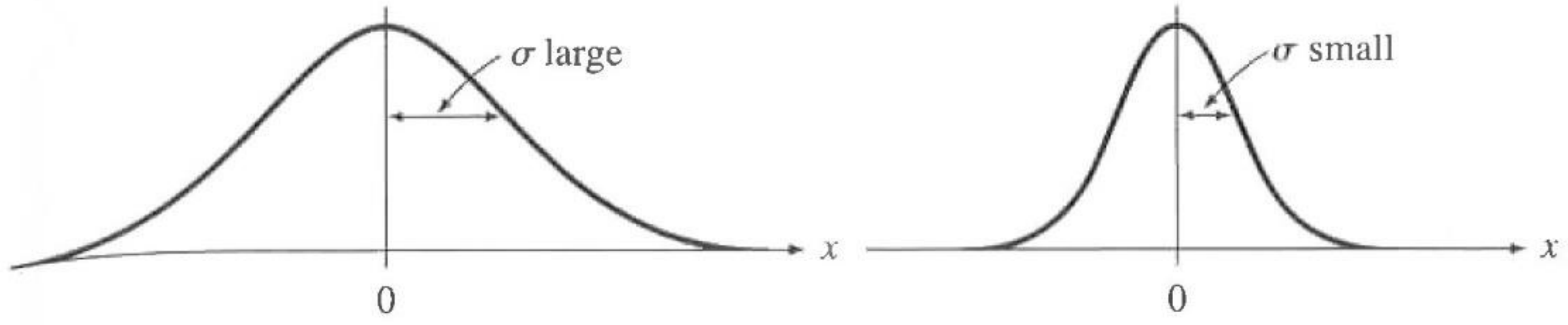
- The error propagation formulas that you will have encountered are actually derived by assuming that measurements have a Gaussian spectrum of values.
- Gaussian distributions appear frequently in physics and other areas of science.

The prototype Gauss function:

$$f(x) = e^{-x^2/2\sigma^2} .$$

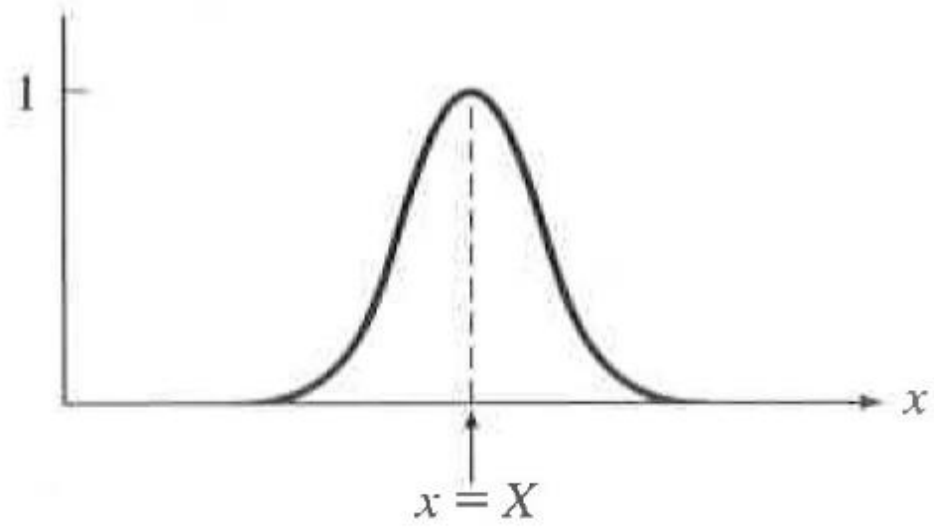
- This function is symmetric about $x = 0$.
- Has the value of $y = 1$ at $x = 0$.
- The parameter σ specifies the width of the distribution.
- Has a non-zero value (i.e. $y > 0$) from $-\infty$ to $+\infty$.
- Once x is more than a few times σ the value of the function becomes negligibly small.

A large value of σ specifies a wider curve.



- Usually the Gaussian is centred on some particular value, X.

i.e., $$f(x) = e^{-(x-X)^2/2\sigma^2}\ .$$



A Gaussian centred on the value $x = X$.

An additional requirement that we must have for a Gaussian to represent the probability of finding a certain measurement: $\Rightarrow$ The area must be normalised to 1.

i.e., the measurement must occur somewhere under the curve.

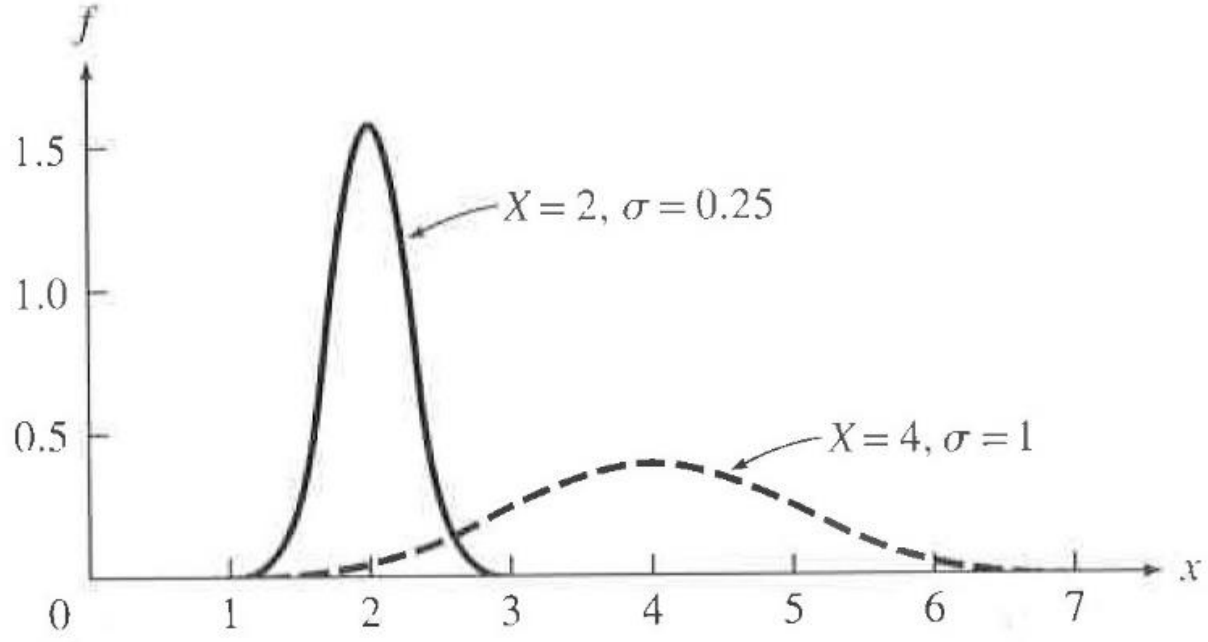
we require: $$\int_{-\infty}^{+\infty} f(x)dx = 1\,.$$

To arrange the normalization: $$f(x) = A\, e^{-\frac{(x-X)^2}{2\sigma^2}}.$$

Choose the value of A so that the integral is satisfied.

Normalised Gaussian function:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-X)^2}{2\sigma^2}}.$$



Two area-normalised Gaussians.

Properties of the normal distribution

Mean value:

- The mean value $\bar{x}$ is just the centre position of the distribution (X).
- This is obvious from the symmetry of the distribution.
- The mean value is also the most probable value.

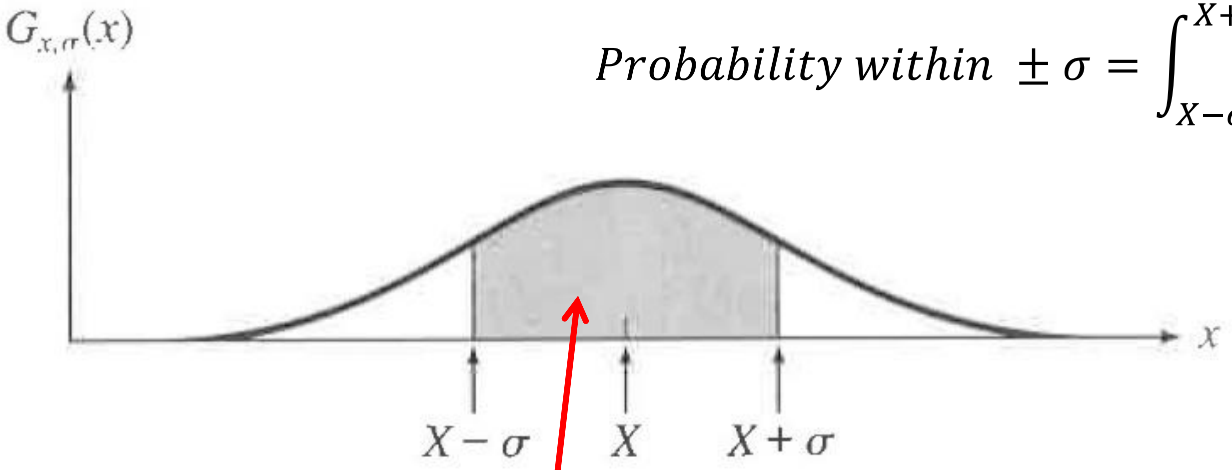
Standard deviation:

- The standard deviation of the distribution is the width parameter, σ.

The standard deviation as the 68% confidence limit

Based on the normal distribution, the probability of any one measurement being with the range $a \leq x \leq b$ is given by:

$$Probability = \int_a^b f(x)dx \ .$$



$$Probability\ within\ \pm\sigma = \int_{X-\sigma}^{X+\sigma} G(x)dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{X-\sigma}^{X+\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}} dx$$

probability of a measurement within one standard deviation of X (shaded).

$$Probability\ within\ \pm\sigma = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-z^2/2} dz\ .$$

No of sigma ($t = 1$ here)

Known as the error function

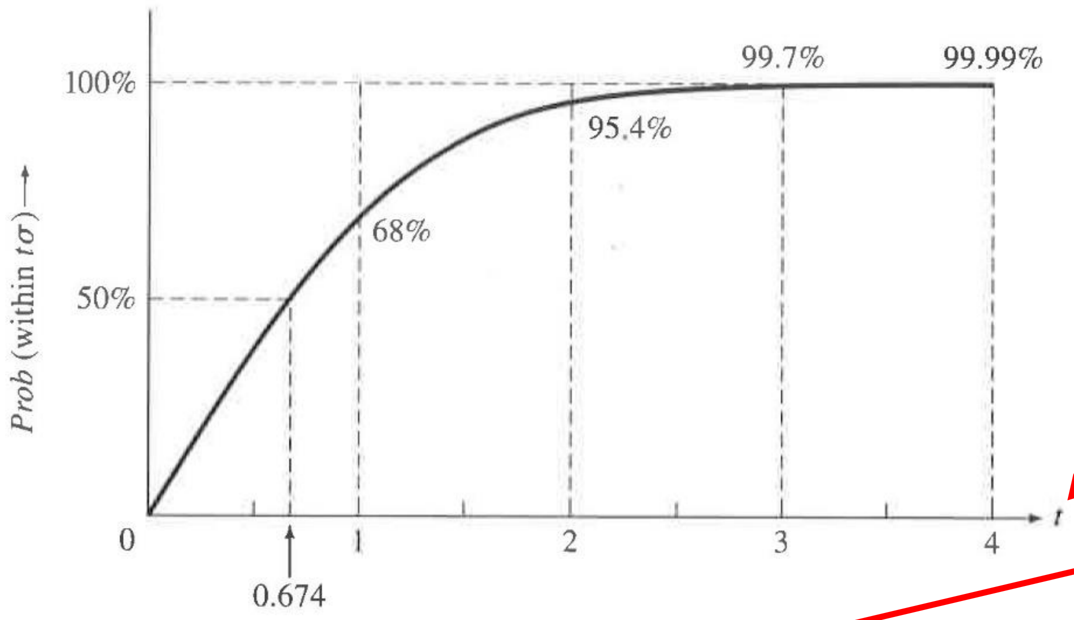
$$= \mathrm{erf}\left(\frac{t}{\sqrt{2}}\right)$$

- erf() has no analytic solution and must be evaluated numerically.

$$Probability\ within\ \pm\sigma = 0.683\ .$$

Consequently it is expected that about 68% of the time, a measurement will be within $\pm 1\ \sigma$.

We can calculate the probability for a measurement to be within any arbitrary value of $t\sigma$, which is graphed below:



No. of std deviations

In physics, measurements are quoted using uncertainties corresponding to the standard deviation (i.e. "1 sigma" uncertainties).

t	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5	4.0
Prob (%)	0	20	38	55	68	79	87	92	95.4	98.8	99.7	99.95	99.99

- Other disciplines (such as humanities) use two-sigma as the standard measure of the uncertainty (95%).

Exception: Discovery in Particle Physics:

- The traditional threshold of certainty for stating that a discovery has been made is that the chance of the background having a statistical fluctuation at least as large as the observed particle is equivalent to a 5-sigma event.
- This corresponds to a probability of around 3×10^{-7} that a random fluctuation could masquerade as the particle (only need to consider $+5\sigma$ probability and not the -5σ here).

Software data fitting:

- Many software in their fitting routines calculate uncertainties in the fitted values.
- The uncertainties usually quoted as the "95% confidence level" or similar wording.
- This means that uncertainty is quoted as a 2-sigma value.
- To convert to 1-sigma values, just halve the 2-sigma value.

For your work in the laboratory, you should always quote uncertainty as the 1-sigma value (the standard deviation).