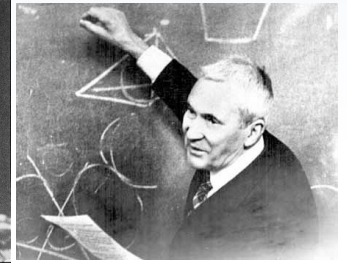
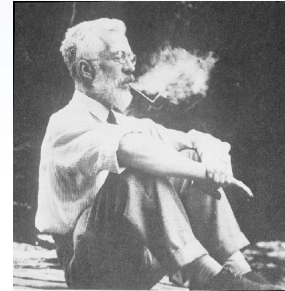
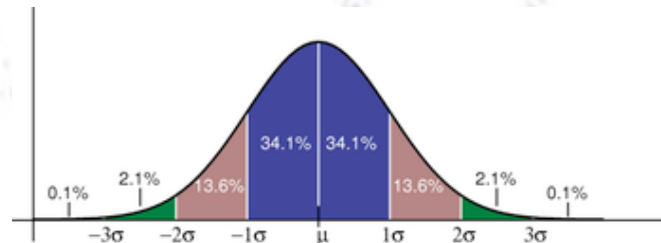


Applied Statistics

Binomial, Poisson, and Gaussian



Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Probability Density Functions

A Probability Density Function (PDF) $f(x)$ describes the probability of an outcome x :

probability to observe x in the interval $[x, x+dx] = f(x) dx$

PDFs are required to be normalised:

$$\int_S f(x) dx = 1$$

The expectation value (aka. mean) and the variance (i.e. standard deviation squared) are then defined as follows:

$$\mu = \int_{-\infty}^{\infty} x f(x) dx$$

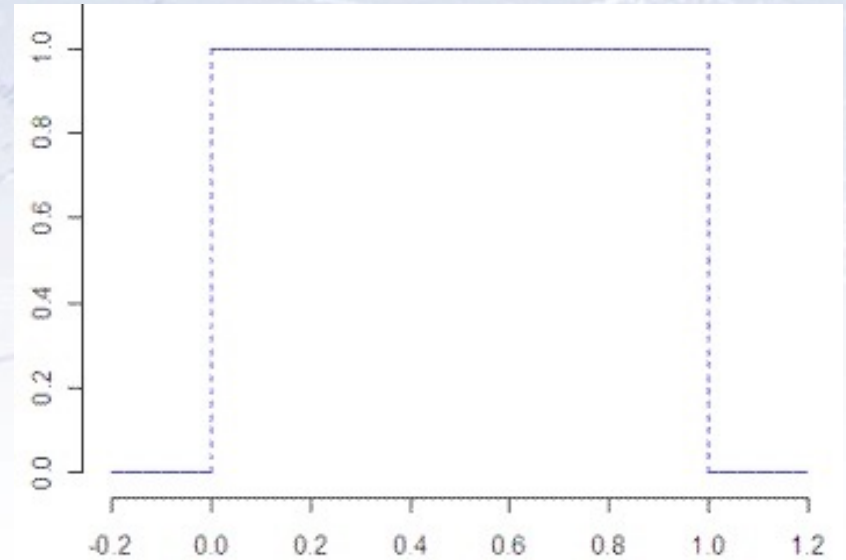
$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

Probability Density Functions

Example:

Consider a uniform distribution:

$$f(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & \textit{else} \end{cases}$$



Calculating the mean and width:

$$\mu = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x dx = \left[\frac{1}{2} x^2 \right]_0^1 = \frac{1}{2}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx =$$

$$\left[\frac{1}{3} x^3 - \frac{1}{2} x^2 + \frac{1}{4} x \right]_0^1 = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$$

Probability Density Functions

The number of PDFs is infinite, and nearly so is the list of known ones:

Discrete distributions [\[edit source | edit beta \]](#)

With finite support [\[edit source | edit beta \]](#)

- The Bernoulli distribution, which takes value 1 with probability p .
- The Rademacher distribution, which takes value ± 1 with equal probability.
- The binomial distribution, which describes the number of successes in a fixed number of independent trials.
- The beta-binomial distribution, which describes the distribution of the number of successes in a fixed number of trials when the probability of success is itself a random variable.
- The degenerate distribution at x_0 , where X is certain to take the value x_0 .
- The discrete uniform distribution, where all elements of a finite set are equally likely.
- The hypergeometric distribution, which describes the number of successes in a fixed number of trials without replacement.
- The Poisson binomial distribution, which describes the number of successes in a fixed number of trials with varying probabilities.
- Fisher's noncentral hypergeometric distribution
- Wallenius' noncentral hypergeometric distribution
- Benford's law, which describes the frequency of digits in many real-life sets of numerical data.

With infinite support [\[edit source | edit beta \]](#)

- The beta negative binomial distribution
- The Boltzmann distribution, a discrete distribution in statistical mechanics. Special cases include:
 - The Gibbs distribution
 - The Maxwell-Boltzmann distribution
- The Borel distribution
- The extended negative binomial distribution
- The extended hypergeometric distribution
- The generalized log-series distribution
- The generalized normal distribution
- The geometric distribution, a discrete distribution with infinite support.
- The hypergeometric distribution
- The logarithmic (series) distribution
- The negative binomial distribution or Pascal distribution
- The parabolic fractal distribution
- The Poisson distribution, which describes a very large number of independent trials with a small probability of success.
- The Conway-Maxwell-Poisson distribution, a two-parameter generalization of the Poisson distribution.
- The Polya-Eggenberger distribution
- The Skellam distribution, the distribution of the difference of two independent random walks.
- The skew elliptical distribution
- The skew normal distribution
- The Yule-Simon distribution
- The zeta distribution has uses in applied statistics
- Zipf's law or the Zipf distribution. A discrete power-law distribution.
- The Zipf-Mandelbrot law is a discrete power-law distribution.

Continuous distributions [\[edit source | edit beta \]](#)

Supported on a bounded interval [\[edit source | edit beta \]](#)

- The Arcsine distribution on $[a, b]$, which is a special case of the beta distribution.
- The Beta distribution on $[0, 1]$, of which the uniform distribution is a special case.
- The Lognormal distribution on $(0, \infty)$.
- The Dirac delta function although not strictly a function, but the notation treats it as if it were a continuous distribution.
- The continuous uniform distribution on $[a, b]$, where:
 - The rectangular distribution is a uniform distribution on $[a, b]$.
- The Irwin-Hall distribution is the distribution of the sum of n independent uniform random variables.
- The Kent distribution on the three-dimensional sphere.
- The Kumaraswamy distribution is as versatile as the beta distribution.
- The logarithmic distribution (continuous)
- The PERT distribution is a special case of the beta distribution.
- The raised cosine distribution on $[\mu - s, \mu + s]$
- The reciprocal distribution
- The triangular distribution on $[a, b]$, a special case of the beta distribution.
- The truncated normal distribution on $[a, b]$.
- The U-quadratic distribution on $[a, b]$.
- The von Mises distribution on the circle.
- The von Mises-Fisher distribution on the N -dimensional sphere.
- The Wigner semicircle distribution is important in quantum mechanics.

Supported on semi-infinite intervals, usually $[0, \infty)$

- The Beta prime distribution
- The Birnbaum-Saunders distribution, also known as the Rayleigh distribution
- The chi distribution
 - The noncentral chi distribution
- The chi-squared distribution, which is the sum of the squares of k independent standard normal random variables.
 - The inverse-chi-squared distribution
 - The noncentral chi-squared distribution
 - The Scaled-inverse-chi-squared distribution
- The Dagum distribution
- The exponential distribution, which describes the time between events in a Poisson process.
- The F-distribution, which is the distribution of the ratio of two chi-squared variates which are not independent.
- The noncentral F-distribution
- Fisher's z-distribution
- The folded normal distribution
- The Fréchet distribution
- The Gamma distribution, which describes the time between events in a Poisson process.
 - The Erlang distribution, which is a special case of the gamma distribution.
 - The inverse-gamma distribution
- The generalized Pareto distribution
- The Gamma/Gompertz distribution
- The Gompertz distribution
- The half-normal distribution

- Hotelling's T-squared distribution
- The inverse Gaussian distribution, also known as the Wald distribution
- The Lévy distribution
- The log-Cauchy distribution
- The log-gamma distribution
- The log-Laplace distribution
- The log-logistic distribution
- The log-normal distribution, describing variability in many natural phenomena
- The Mittag-Leffler distribution
- The Nakagami distribution
- The Pareto distribution, or "power law" distribution
- The Pearson Type III distribution
- The phased bi-exponential distribution is common in pharmacokinetics
- The phased bi-Weibull distribution
- The Rayleigh distribution
- The Rayleigh mixture distribution
- The Rice distribution
- The shifted Gompertz distribution
- The type-2 Gumbel distribution
- The Weibull distribution or Rosin-Rammler distribution, used in grinding, milling and crushing operations.

Supported on the whole real line [\[edit source | edit beta \]](#)

- The Behrens-Fisher distribution, which arises in the analysis of variance
- The Cauchy distribution, an example of a heavy-tailed distribution
- The resonance energy distribution, impact and fracture
- Chernoff's distribution
- The Exponentially modified Gaussian distribution
- The Fisher-Tippett, extreme value, or log-normal distribution
 - The Gumbel distribution, a special case of the Fisher-Tippett distribution
- Fisher's z-distribution
- The generalized logistic distribution
- The generalized normal distribution
- The geometric stable distribution
- The Holtmark distribution, an example of a heavy-tailed distribution
- The hyperbolic distribution
- The hyperbolic secant distribution
- The Johnson SU distribution
- The Landau distribution
- The Laplace distribution
- The Lévy skew alpha-stable distribution or stable distribution, Lévy distribution and normal distribution
- The Linnik distribution
- The logistic distribution
- The map-Airy distribution
- The normal distribution, also called the Gaussian distribution, independent, identically distributed variables
- The Normal-exponential-gamma distribution
- The Pearson Type IV distribution (see Pearson distribution)
- The skew normal distribution

- Student's t-distribution, useful for estimating the mean of a normal distribution with unknown variance.
 - The noncentral t-distribution
- The type-1 Gumbel distribution
- The Voigt distribution, or Voigt profile, is the convolution of a Gaussian and a Lorentzian distribution.
- The Gaussian minus exponential distribution is used in queueing theory.

With variable support [\[edit source | edit beta \]](#)

- The generalized extreme value distribution has a shape parameter
- The generalized Pareto distribution has a shape parameter
- The Tukey lambda distribution is either supported on a bounded interval or on the whole real line
- The Wakeby distribution

Mixed discrete/continuous distributions [\[edit source | edit beta \]](#)

- The rectified Gaussian distribution replaces the Gaussian distribution with a uniform distribution on the positive real line.

Joint distributions [\[edit source | edit beta \]](#)

For any set of independent random variables the joint distribution is the product of the individual distributions.

Two or more random variables on the same space

- The Dirichlet distribution, a generalization of the multinomial distribution
- The Ewens's sampling formula is a probability distribution on permutations
- The Balding-Nichols model
- The multinomial distribution, a generalization of the binomial distribution
- The multivariate normal distribution, a generalization of the normal distribution
- The negative multinomial distribution, a generalization of the negative binomial distribution
- The generalized multivariate log-gamma distribution

Matrix-valued distributions [\[edit source | edit beta \]](#)

- The Wishart distribution
- The inverse-Wishart distribution
- The matrix normal distribution
- The matrix t-distribution

Non-numeric distributions [\[edit source | edit beta \]](#)

- The categorical distribution
- newton distribution

Miscellaneous distributions [\[edit source | edit beta \]](#)

- The Cantor distribution
- The generalized logistic distribution family
- The Pearson distribution family
- The phase-type distribution

And surely more!

Probability Density Functions

The number of PDFs is infinite, and nearly so is the list of known ones:

Discrete distributions [edit source | edit beta]

With finite support [edit source | edit beta]

- The Bernoulli distribution, which takes value 1 with
- The Rademacher distribution, which takes value 1
- The binomial distribution, which describes the num
- The beta-binomial distribution, which describes the
- The degenerate distribution at x_0 , where X is certa
- The discrete uniform distribution, where all element
- The shuffled deck

- The
- there
- The f
- Fisher
- Waller
- Benf

With infi

- The t
- The E
- anal
- Th
- Th
- The E
- The e
- The generalized log-series distribution
- The generalized normal distribution
- The geometric distribution, a discrete distribution w
- The hypergeometric distribution
- The logarithmic (series) distribution
- The negative binomial distribution or Pascal distrib
- The parabolic fractal distribution
- The Poisson distribution, which describes a very la
- The Poisson, the hyper-Poisson, the general Poisson t
- The Conway–Maxwell–Poisson distribution, a tw
- The Polya–Eggenberger distribution
- The Skellam distribution, the distribution of the diff
- The skew elliptical distribution
- The skew normal distribution
- The Yule–Simon distribution
- The zeta distribution has uses in applied statistics
- Zipf's law or the Zipf distribution. A discrete power-
- The Zipf–Mandelbrot law is a discrete power law dis

Continuous distributions [edit source | edit beta]

Supported on a bounded interval [edit source | edit

- The Arcsine distribution on $[a,b]$, which is a speci
- The Beta distribution on $[0,1]$, of which the unifor
- The Logitnormal distribution on $(0,1)$.
- The Dirac delta function although not strictly a fur
- The continuous uniform distribution on $[a,b]$, when
- The rectangular distribution is a uniform distrib
- The Irwin–Hall distribution is the distribution of the

- Hotelling's T-squared distribution
- The inverse Gaussian distribution, also kn
- The Lévy distribution
- The log-Cauchy distribution
- The log-gamma distribution
- The log-Laplace distribution
- The log-logistic distribution
- The log-normal distribution, describing vari
- The Mittag-Leffler distribution
- The Nakagami distribution
- The Pareto distribution, or "power law" dist
- The Pearson Type III distribution

- Student's t-distribution, useful for estimating t
 - The noncentral t-distribution
 - The type-1 Gumbel distribution
 - The Voigt distribution, or Voigt profile, is the c
 - The Gaussian minus exponential distribution is
- With variable support [edit source | edit beta]
- The generalized extreme value distribution has
 - parameter
 - The generalized Pareto distribution has a supp
 - The Tukey lambda distribution is either suppor
 - The Wakeby distribution

ONS [ed
replaces n
beta]
ables the
same sar
zation of t
probability
ralization c
a general
h, a gener
ma distrib
nce | edit t

"Essentially, all models are wrong, but some are useful"

[George E. P. Box, British Statistician, 1919-2013]

- The chi-squared distribution, which is the sum of t
- The inverse-chi-squared distribution
- The noncentral chi-squared distribution
- The Scaled-inverse-chi-squared distribution
- The Dagum distribution
- The exponential distribution, which describes the t
- The F-distribution, which is the distribution of the r
- ratio of two chi-squared variates which are not no
- The noncentral F-distribution
- Fisher's z-distribution
- The folded normal distribution
- The Fréchet distribution
- The Gamma distribution, which describes the time
- The Erlang distribution, which is a special cas
- The inverse-gamma distribution
- The generalized Pareto distribution
- The Gamma/Gompertz distribution
- The Gompertz distribution
- The half-normal distribution

- Fisher's z-distribution
- The generalized logistic distribution
- The generalized normal distribution
- The geometric stable distribution
- The Holtzmark distribution, an example of
- The hyperbolic distribution
- The hyperbolic secant distribution
- The Johnson SU distribution
- The Landau distribution
- The Laplace distribution
- The Lévy skew alpha-stable distribution or
- distribution, Lévy distribution and normal c
- The Linnik distribution
- The logistic distribution
- The map-Airy distribution
- The normal distribution, also called the Ga
- independent, identically distributed variabl
- The Normal-exponential-gamma distributor
- The Pearson Type IV distribution (see Pea
- The skew normal distribution

- The matrix t-distribution
- Non-numeric distributions [edit source | edit t
- The categorical distribution
 - newton distribution
- Miscellaneous distributions [edit source | edit
- The Cantor distribution
 - The generalized logistic distribution family
 - The Pearson distribution family
 - The phase-type distribution

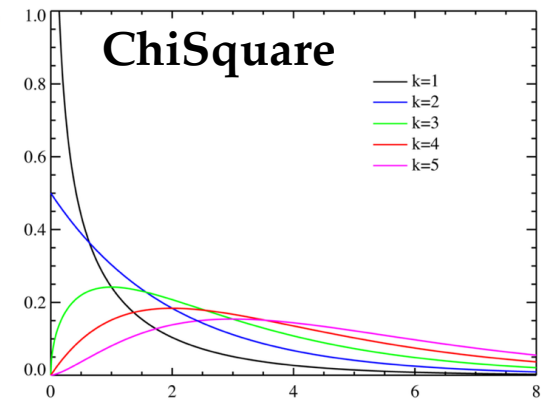
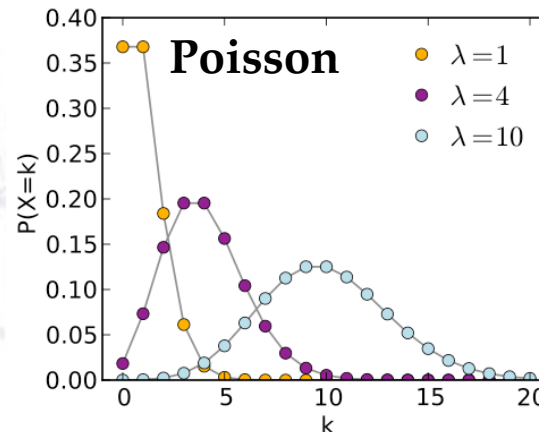
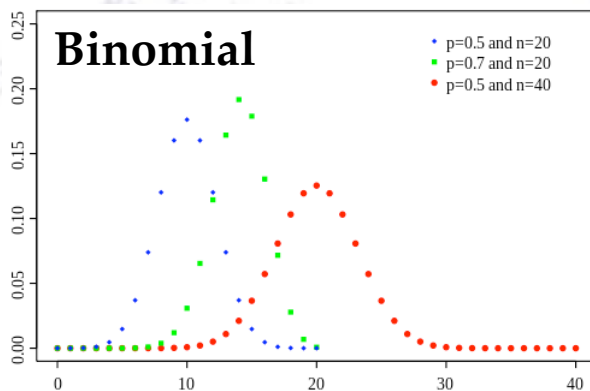
And surely more!

Probability Density Functions

An almost complete list of those we will deal with in this course is:

- **Gaussian** (aka. Normal)
- **Poisson**
- **Binomial** (and also Multinomial)
- Students t-distribution
- Uniform
- ChiSquare
- Exponential
- Error function (integral of Gaussian)

You should already know most of these, and the rest will be explained.



Binomial, Poisson, Gaussian

$$f(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

Given **N** trials each with **p** chance of **success**, how many **successes n** should you expect in total?

Note: $n \in \mathbb{N}$, $p \in [0, 1]$

This distribution is... **Binomial**:

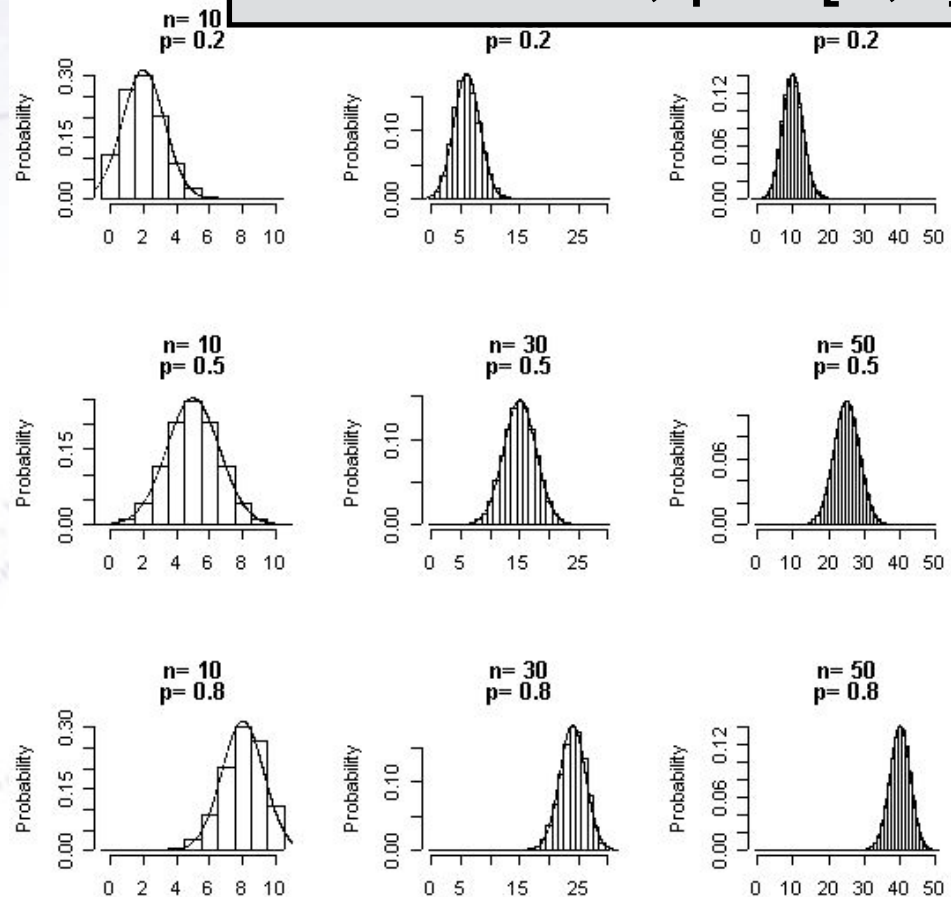
$$f(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

Mean = Np

Variance = $Np(1-p)$

This means, that the error on a fraction $f = n/N$ is:

$$\sigma(f) = \sqrt{\frac{f(1-f)}{N}}$$



Binomial, Poisson, Gaussian

You count 100 cars, and see 15 red cars. What is your estimate of the fraction (i.e. probability) of red cars and its uncertainty?

- a) 0.150 ± 0.050
- b) 0.150 ± 0.026
- c) 0.150 ± 0.036
- d) 0.125 ± 0.030
- e) 0.150 ± 0.081

From previous page:
$$\sigma(f) = \sqrt{\frac{f(1-f)}{N}}$$

A friend tells you, that 8% of the cars on Blegdamsvej are red. What is the chance of that? Could he be right?

Binomial, Poisson, Gaussian

You count 100 cars, and see 15 red cars. What is your estimate of the fraction (i.e. probability) of red cars and its uncertainty?

- a) 0.150 ± 0.050
- b) 0.150 ± 0.026
- c) 0.150 ± 0.036
- d) 0.125 ± 0.030
- e) 0.150 ± 0.081

$$(0.150 - 0.080) / 0.036 = 1.9 \sigma$$

From previous page:
$$\sigma(f) = \sqrt{\frac{f(1-f)}{N}}$$

A friend tells you, that 8% of the cars on Blegdamsvej are red. What is the chance of that? Could he be right?

Binomial, Poisson, Gaussian

Requirements to be Binomial:

- Fixed number of trials, N
- Independent trials.
- Only two outcomes (success/failure).
- Constant probability of success/failure.

If number of possible outcomes is more than two \Rightarrow **Multinomial distribution.**

Examples of Binomial experiments:

- Tossing a coin 20 times counting number of tails.
- Asking 200 people if they watch sports on TV.
- Rolling a die to see if a 6 appear (Multinomial for all outcomes!).
- Asking 100 die-hards from Ehedslisten, if they would vote for Konservative at the next election!

Examples which aren't Binomial experiments:

- Rolling a die until a 6 appears (not fixed number of trials).
- Drawing 5 cards for a poker hand (no replacement \Rightarrow not independent)

Binomial, Poisson, Gaussian

If $N \rightarrow \infty$ and $p \rightarrow 0$, but $Np \rightarrow \lambda$ then a Binomial approaches a Poisson:

$$f(n, \lambda) = \frac{\lambda^n}{n!} e^{-\lambda}$$

In reality, the approximation is already quite good at e.g. $N=50$ and $p=0.1$.

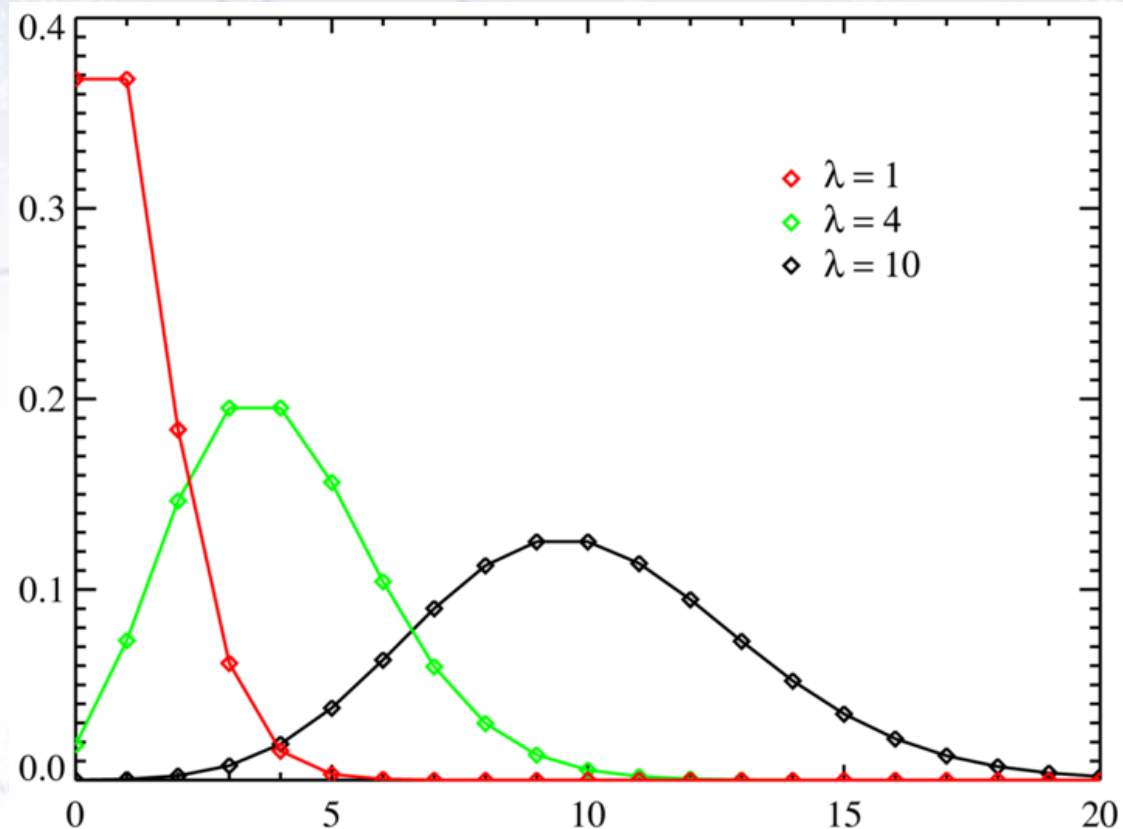
The Poisson distribution only has one parameter, namely λ .

Mean = λ

Variance = λ

So the error on a number is...

...the square root of that number!



Binomial, Poisson, Gaussian

The error on a
(Poisson) number...
is the square root
of that number!!!

The error on a
(Poisson) number...
is the square root
of that number!!!

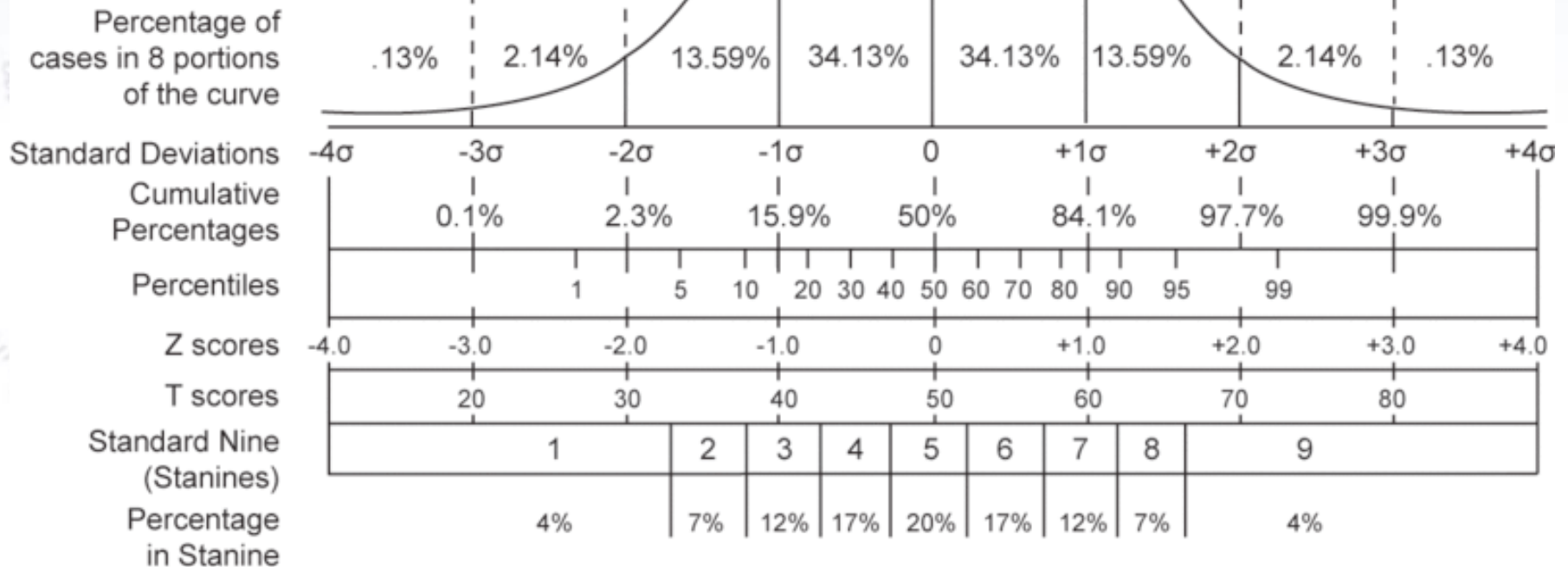
Note: The sum of two Poissons with λ_a and λ_b is a new Poisson with $\lambda = \lambda_a + \lambda_b$.
(See Barlow pages 33-34 for proof)

Binomial, Poisson, Gaussian

If $\lambda \rightarrow \infty$ and many other PDF-cases...

...and for λ , 20 is enough!

*Normal,
Bell-shaped Curve*

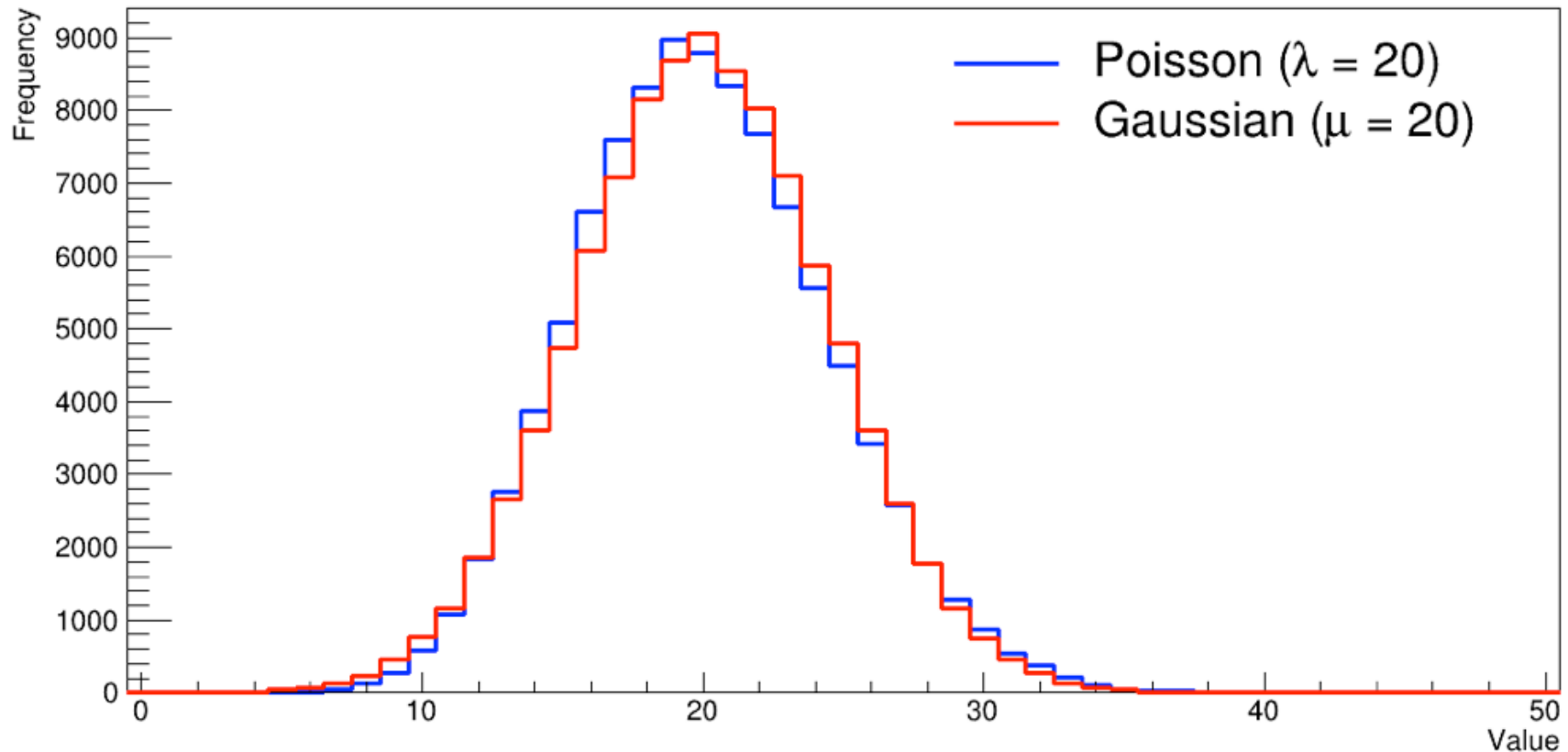


Binomial, Poisson, Gaussian

If $\lambda \rightarrow \infty$ and many other PDF-cases...

...and for λ , 20 is enough!

Poisson and Gaussian distribution comparison

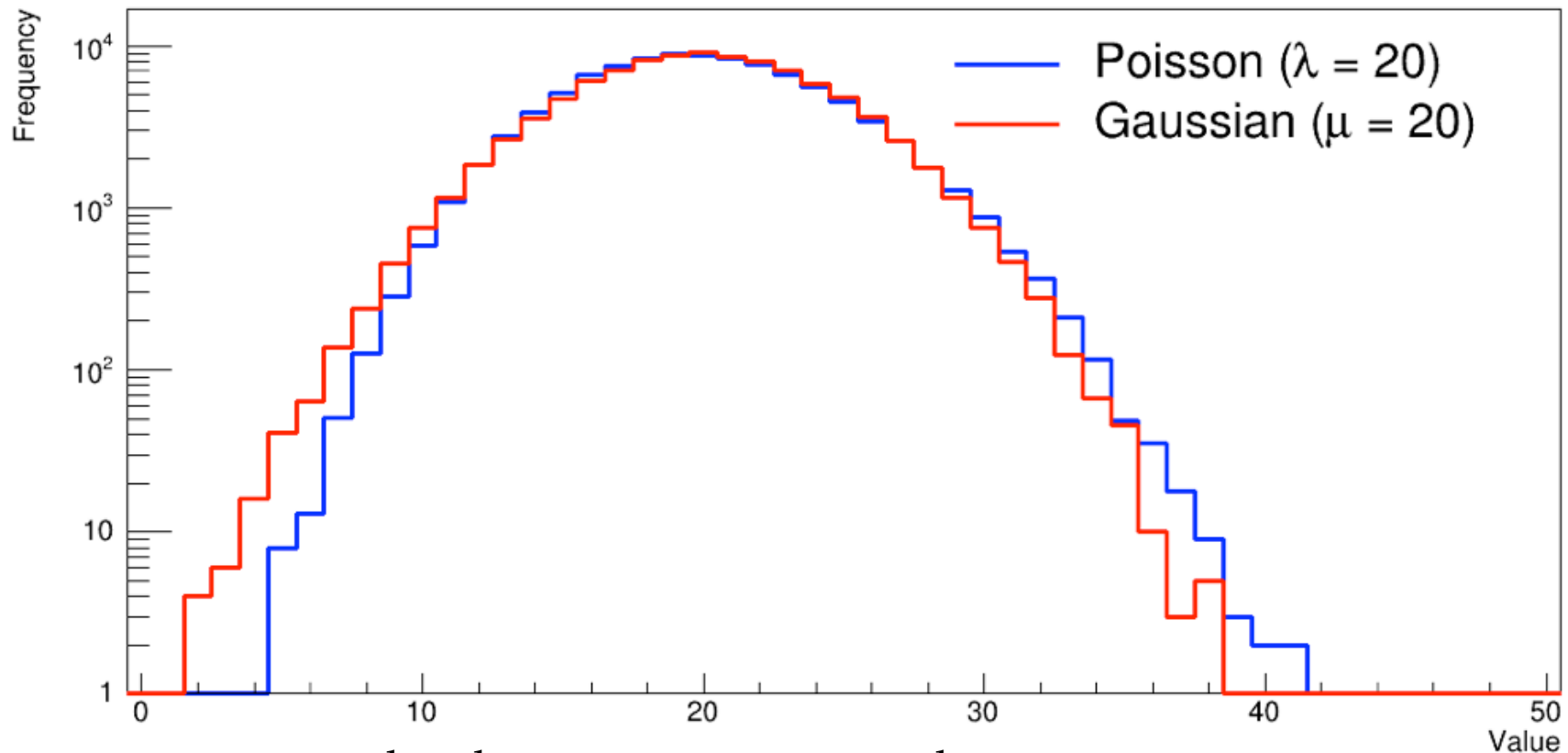


Binomial, Poisson, Gaussian

If $\lambda \rightarrow \infty$ and many other PDF-cases...

...and for λ , 20 is enough!

Poisson and Gaussian distribution comparison

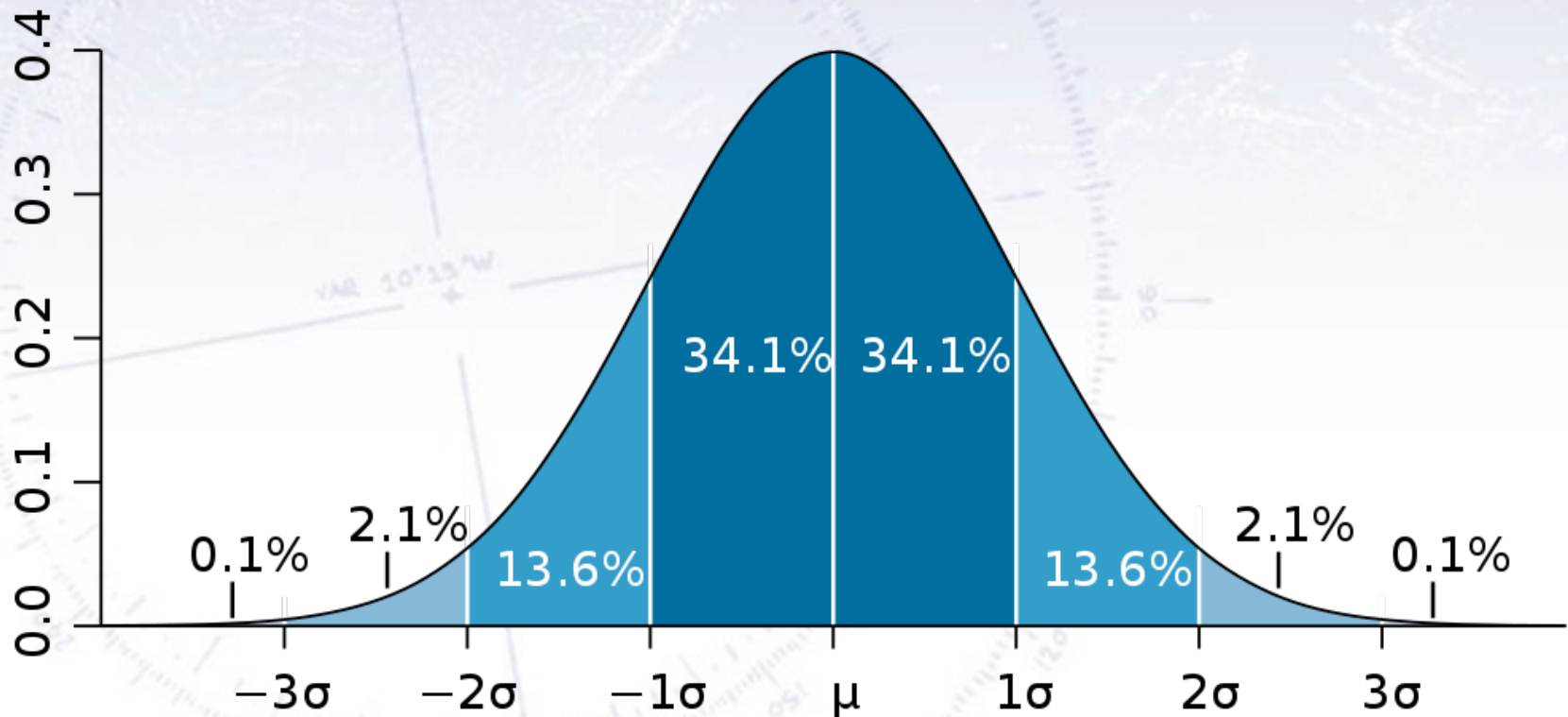


However, note that the TAILS are not quite the same!!!

This is the very reason for the difference between Chi2 and (binned) likelihood!

Binomial, Poisson, Gaussian

“If the Greeks had known it, they would have deified it.”

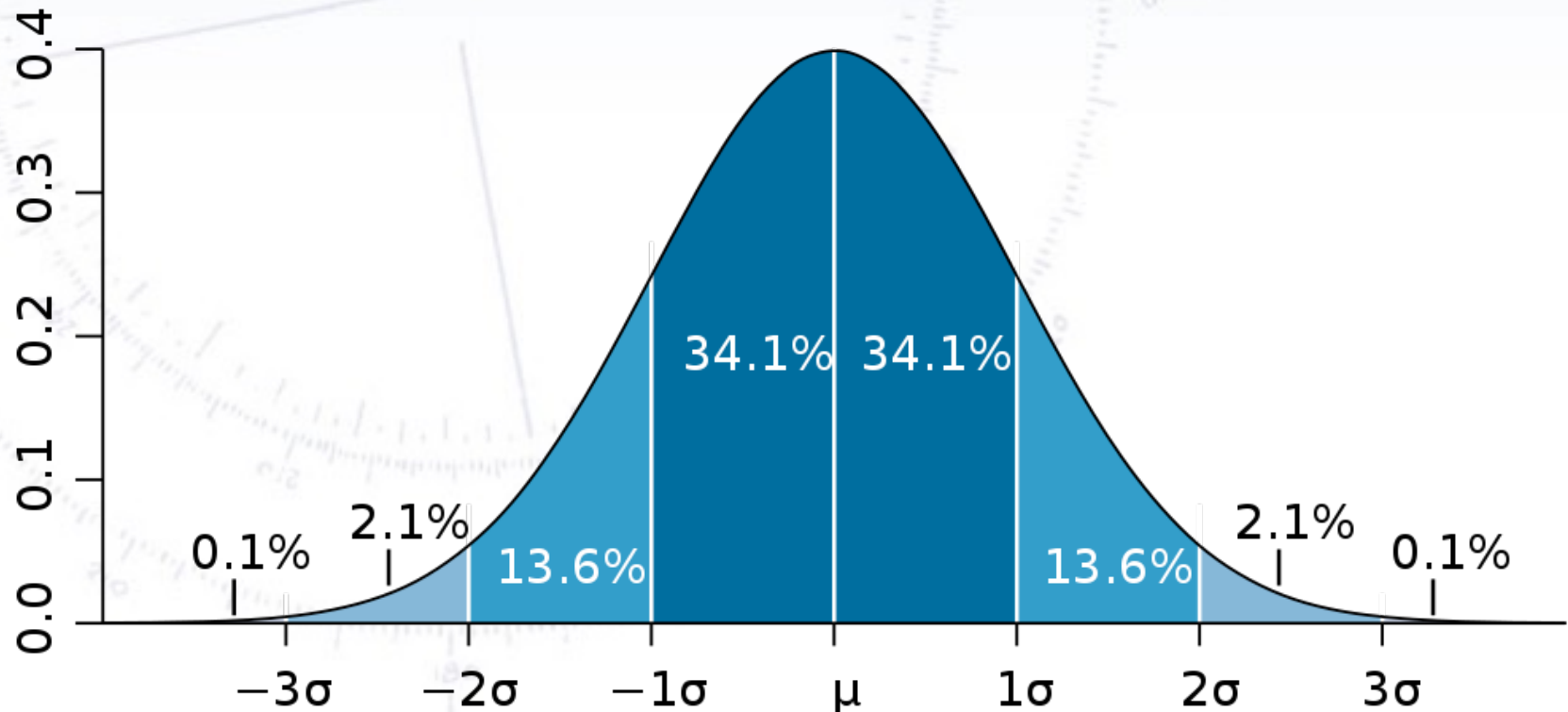


“If the Greeks had known it, they would have deified it. It reigns with serenity and in complete self-effacement amidst the wildest confusion. The more huge the mob and the greater the apparent anarchy, the more perfect is its sway. It is the supreme Law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to be latent all along.” [Karl Pearson]

Binomial, Poisson, Gaussian

The Gaussian defines the way we consider uncertainties.

| Range | Inside | Outside |
|---------------|---------------|-----------|
| $\pm 1\sigma$ | 68 % | 32 % |
| $\pm 2\sigma$ | 95 % | 5 % |
| $\pm 3\sigma$ | 99.7 % | 0.3 % |
| $\pm 5\sigma$ | 99.99995 % | 0.00005 % |



Error function

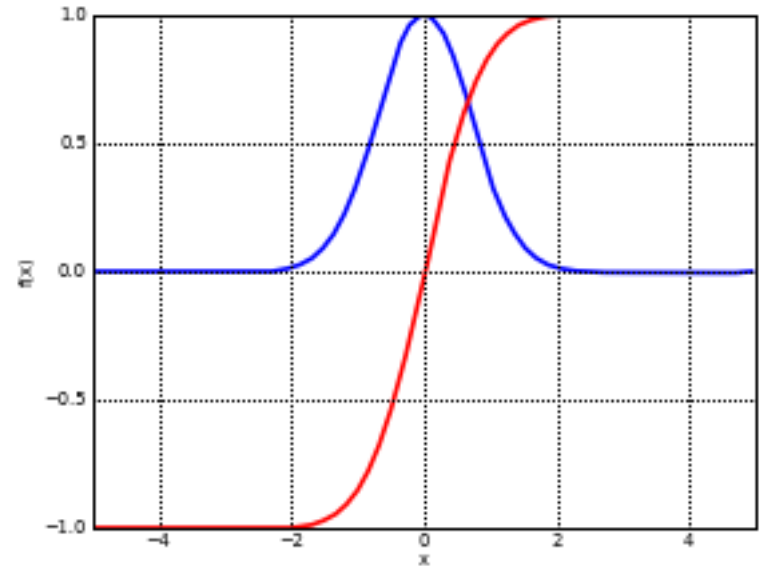
Imagine taking the integral of a Gaussian:

The error function is “almost” that, only it is defined slightly differently, namely as:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

Likewise, there is a complementary error function, which is 1 minus the error function. The functions are used to evaluate Gaussian integrals, i.e. typically “how many sigmas” or “what p-value” does this correspond to.

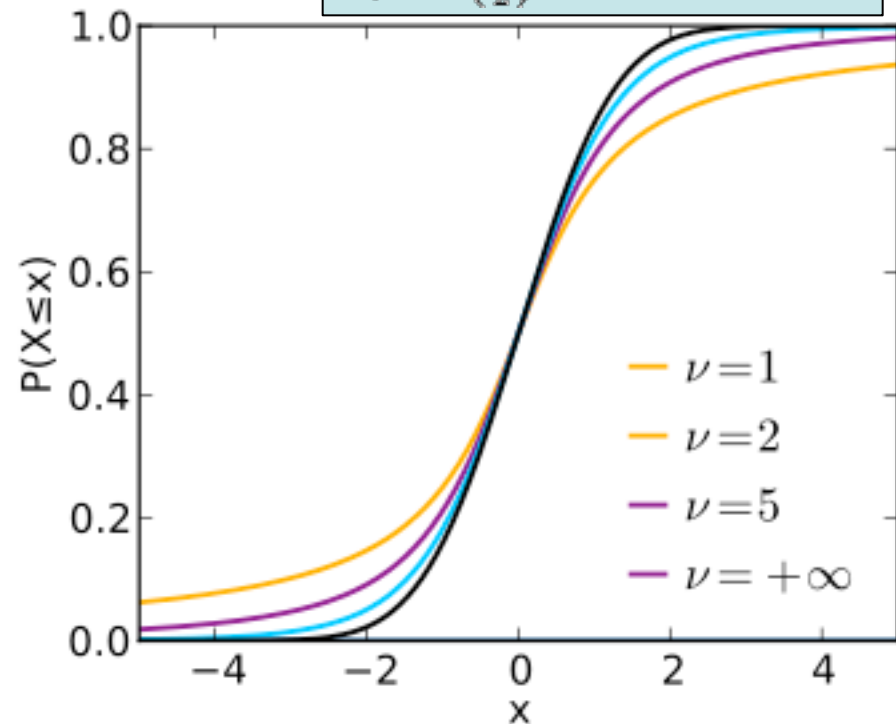
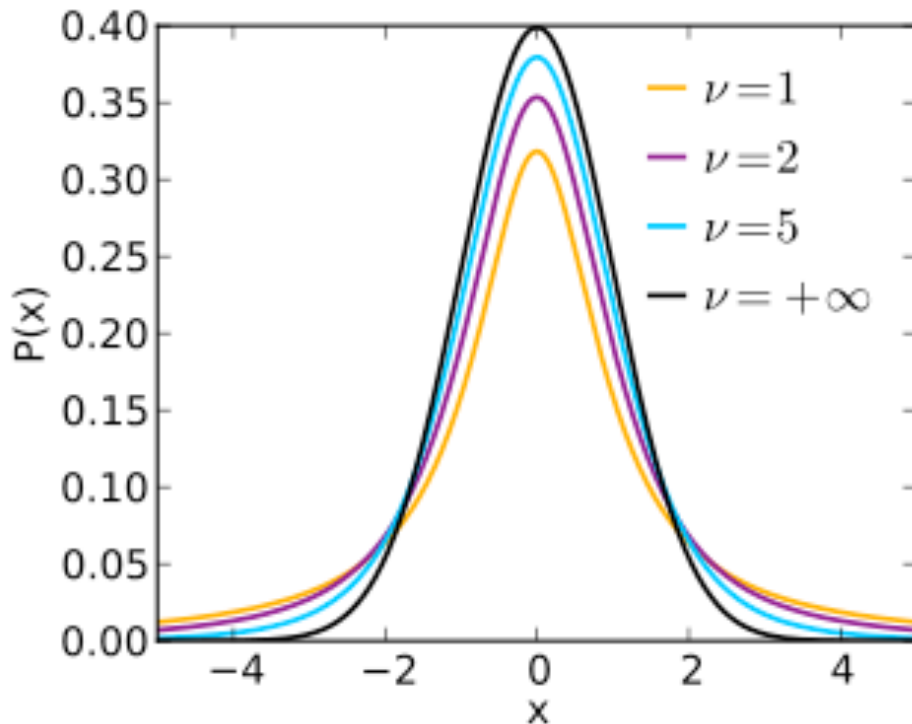
They also belong to the general class of “sigmoids”, i.e. onset functions.



Student's t-distribution

Discovered by William Gosset (who signed "student"), student's t-distribution takes into account **lacking knowledge of the variance**.

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$



When variance is unknown, estimating it from sample gives additional error:

Gaussian:

$$z = \frac{x - \mu}{\sigma}$$

Student's:

$$t = \frac{x - \mu}{\hat{\sigma}}$$