PHYS 575/675

Modern Detectors

Introduction to Statistics and Error Analysis

Outline

- Introduction
- Treatment of experimental data
- Probability distributions
- Maximum likelihood
- Parameter estimation
- Method of least squares
- Bayesian approach
- Hypothesis and significance testing
- Intervals and limits

What and when we need to know about errors? Everyday.

Washington, DC (20064) Weather Save This Location

Updated: Jan 24, 2012, 9:45am EST

SOCIAL NEW! See what people are saying about the weather in Washington

Today | Hourly | Tomorrow | Weekend | 5 day | 10 day | Monthly | Map

View 24-Hour Storm Report for Washington Area

	Today	Tonight	Tomorrow
Cloudy	Partly Cloudy	Clear	Mostly Sunny
7 °C	**12 °**	**-1 °**	**9 °**
Feels Like: 7° Get FREE weather on your desktop	High	Low	High
Past 24-hr: Precip: 1.5 mm (est.) Snow: 0 cm	Chance of Rain: 0%	Chance of Precip: 0%	Chance of Rain: 0%
Wind: From WSW at 5km/h	Wind: W at 11 km/h	Wind: NW at 10 km/h	Wind: NW at 11 km/h
Through 4pm: Partly to mostly sunny with temperatures rising towards the low 50s. Winds W at 5 to 10 mph.	Partly cloudy. High 53F. Winds W at 5 to 10 mph.	Mainly clear skies. Low 31F. Winds NW at 5 to 10 mph.	Except for a few afternoon clouds, mainly sunny. High 48F. Winds NW at 5 to 10 mph.
Humidity: 85%	64%	71%	58%
Dew Point: 5°			
Pressure: 1018.6 mb			
Visibility: 9.7 km			
UV Index: 0 - Low	2 - Low	--	2 - Low

T=7°C± **Best guess**

ΔT~0.5°C

Wind speed 11 km/hr **Best guess**

±1km/hr

What and when we need to know about errors? Science.

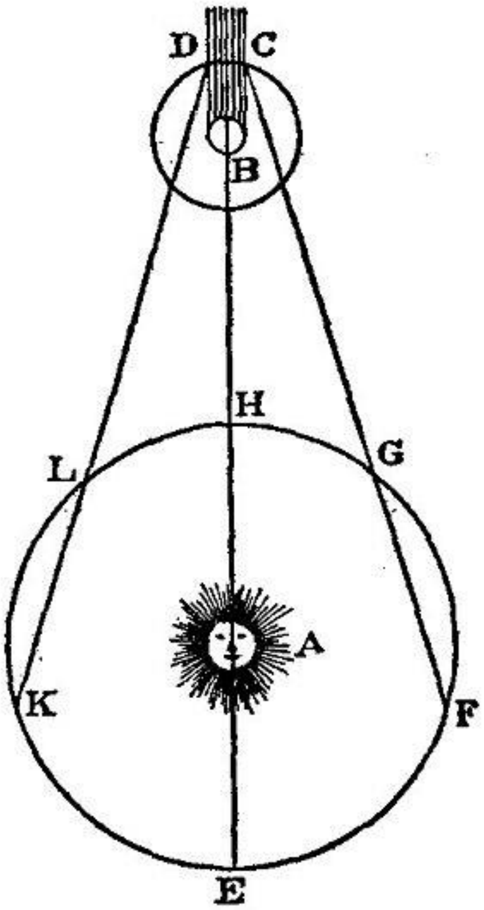


Fig. 70.

From 1676 article on measurement of the speed of light

Measurements of the speed of light

1675 Ole Roemer:

220,000 km/s

NIST Boulder Colorado:

c=299,792,456.2±1.1 m/s

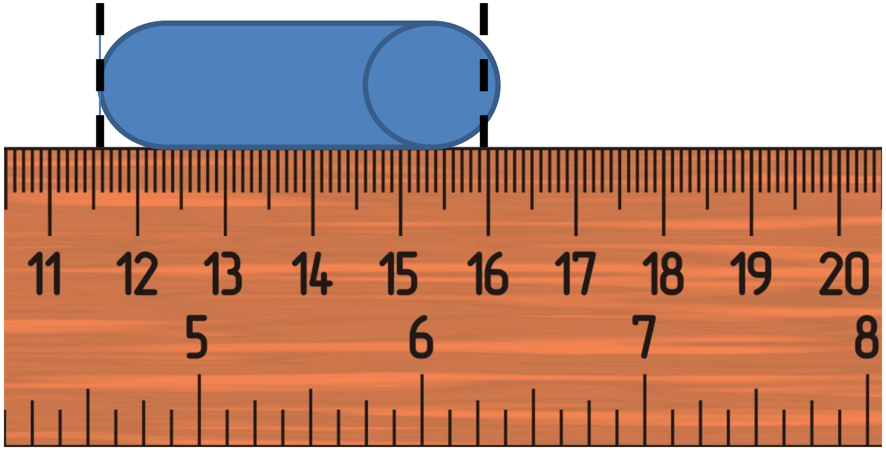
Ole Christensen Romer 1644-1710

Does this make sense?

Something missing?

Reading Error

$\Delta L \approx 0.5$mm **$L = 45$mm$\pm\Delta L$(?)**



$\Delta L \approx 0.03$mm

Requirements on reading error

- Accuracy of better than 1mm not needed
- Ruler is NOT ok, we need to use digital caliper
- Natural limit of accuracy can be due to length uncertainty due to temperature expansion. For 45mm $\Delta L \sim 0.012$mm/K

Reading Error. Digital meters.

Fluke 8845A/8846A Multimeter

Example Vdc (reading)=0.85V

$$\Delta V = 0.85 \cdot (1.8 \text{ x } 10^{-5}) + 1.0 \cdot (0.7 \text{ x } 10^{-5})$$

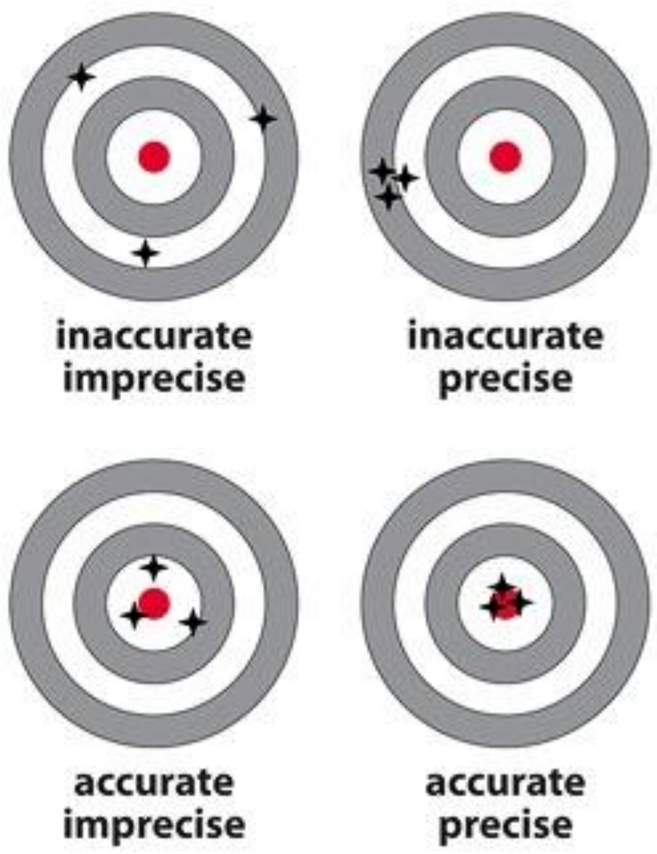
$$\approx 2.2 \text{ x} 10^{-5} = 22\ \mu V$$

8846A Accuracy

Accuracy is given as ± (% measurement + % of range)

Range	24 Hour (23 ±1 °C)	90 Days (23 ±5 °C)	1 Year (23 ±5 °C)	Temperature Coefficient/ °C Outside 18 to 28 °C
100 mV	0.0025 + 0.003	0.0025 + 0.0035	0.0037 + 0.0035	0.0005 + 0.0005
1 V	0.0018 + 0.0006	0.0018 + 0.0007	0.0025 + 0.0007	0.0005 + 0.0001
10 V	0.0013 + 0.0004	0.0018 + 0.0005	0.0024 + 0.0005	0.0005 + 0.0001
100 V	0.0018 + 0.0006	0.0027 + 0.0006	0.0038 + 0.0006	0.0005 + 0.0001
1000 V	0.0018 + 0.0006	0.0031 + 0.001	0.0041 + 0.001	0.0005 + 0.0001

Accuracy vs. Precision



Accuracy of an experiment is measure of how close the results are to the true value

Precision refers to how well individual measurements agree with each other

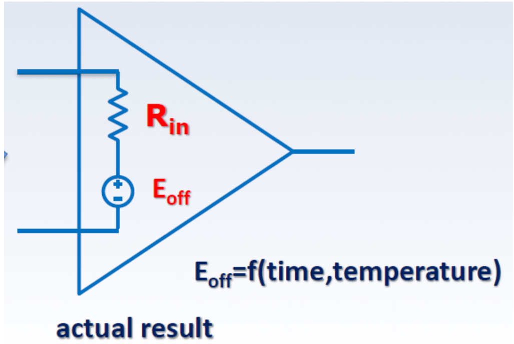
Systematic and Random Errors

Systematic Error: reproducible inaccuracy introduced by faulty equipment, calibration, or technique

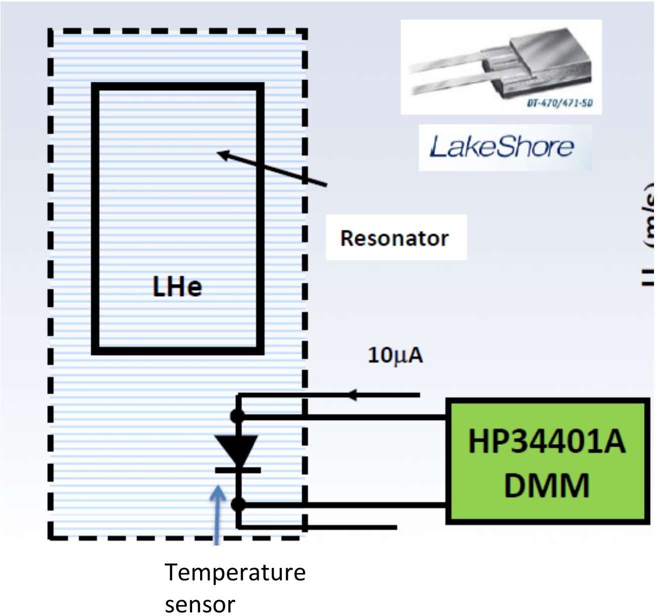
Random Error: indefiniteness of results due to finite precision of experiment. Measure of fluctuation in result after experiment repeated multiple times.

Systematic errors

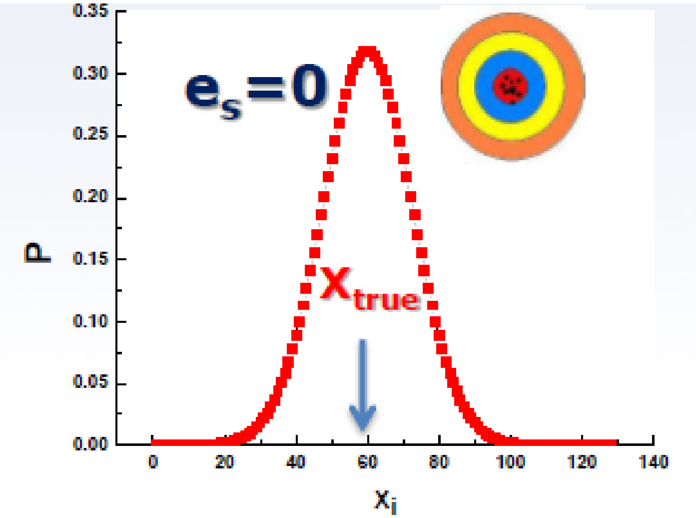
- Sources of systematic errors:
  - Poor calibration of equipment
  - Changes of environmental conditions
  - Imperfect method of observation
  - Drift and offsets in readings

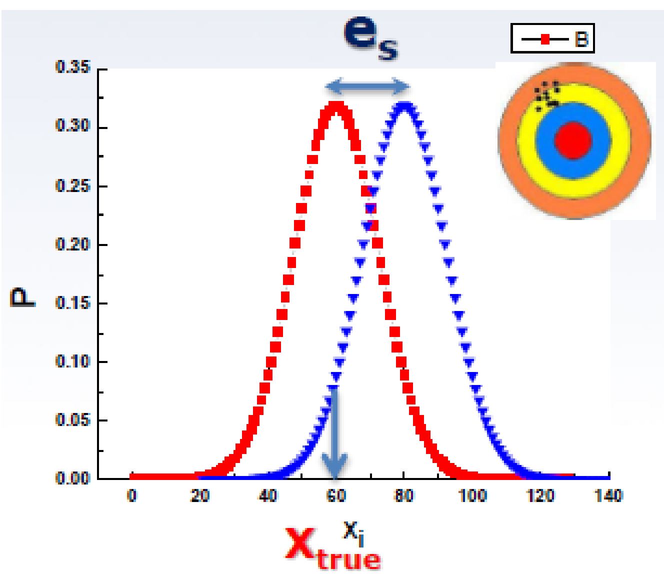


- Examples:
  - Measuring of DC voltage
  - Measuring of polarization current
  - Poor calibration



Random errors





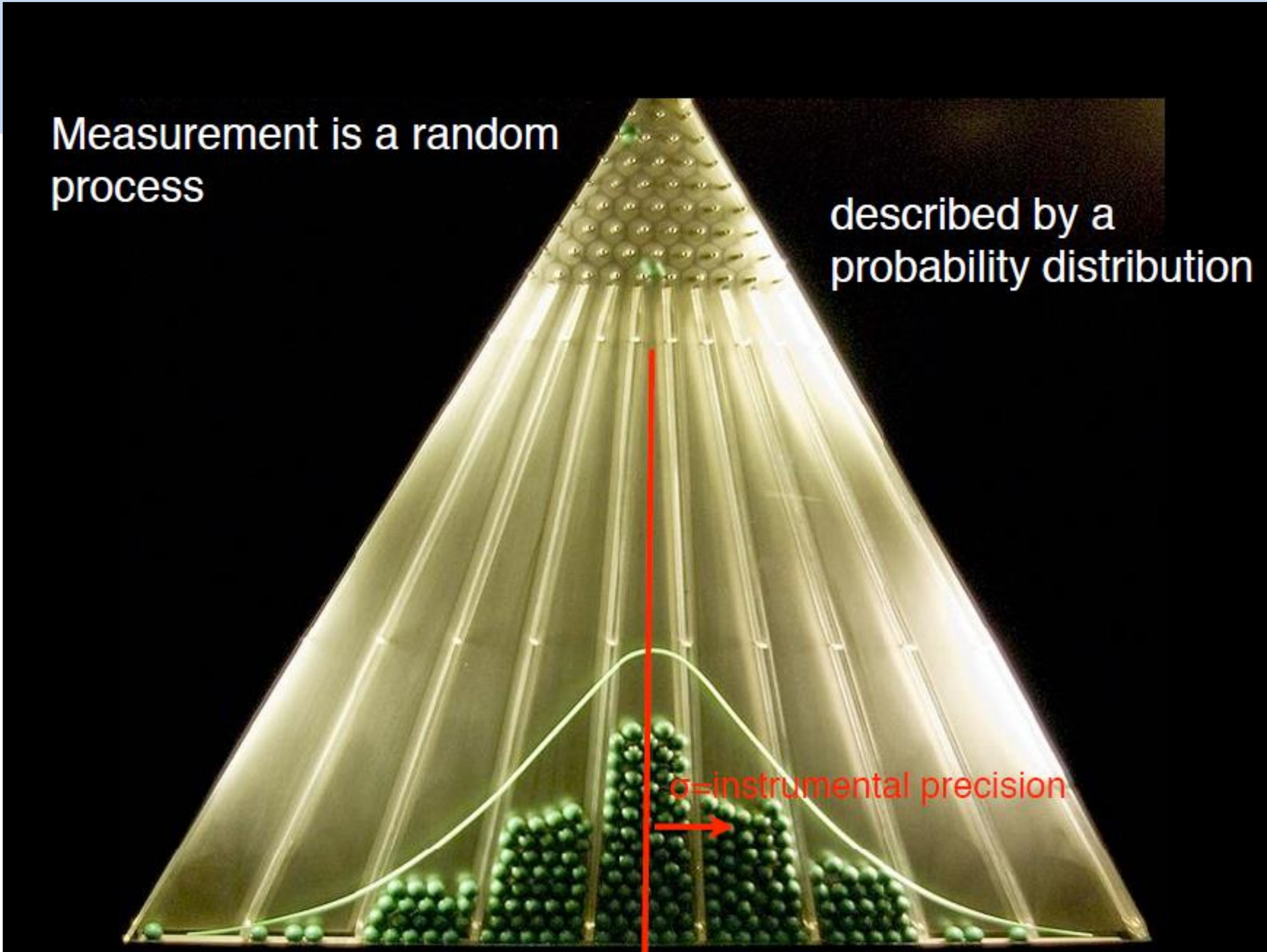
Measured value

systematic error

$$X_{meas} = X_{true} + e_s + e_r$$

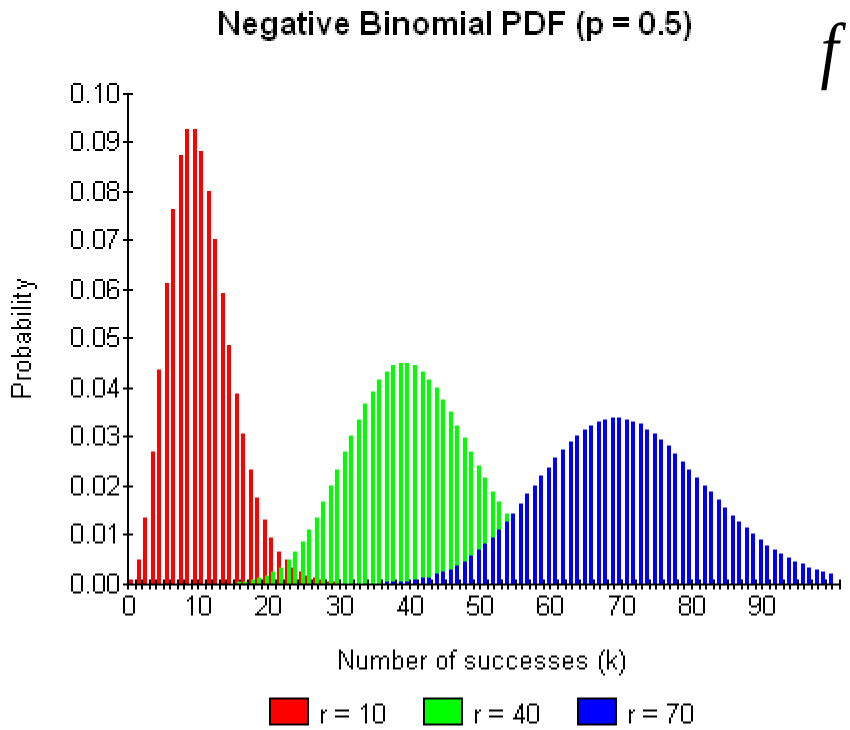
Central value

Random error


Measurement is a random process
described by a probability distribution
σ=instrumental precision

Binomial distribution

For *k+r* Bernoulli trials each with success fraction p, the negative binomial distribution gives the probability of observing k failures and r successes with success on the last trial

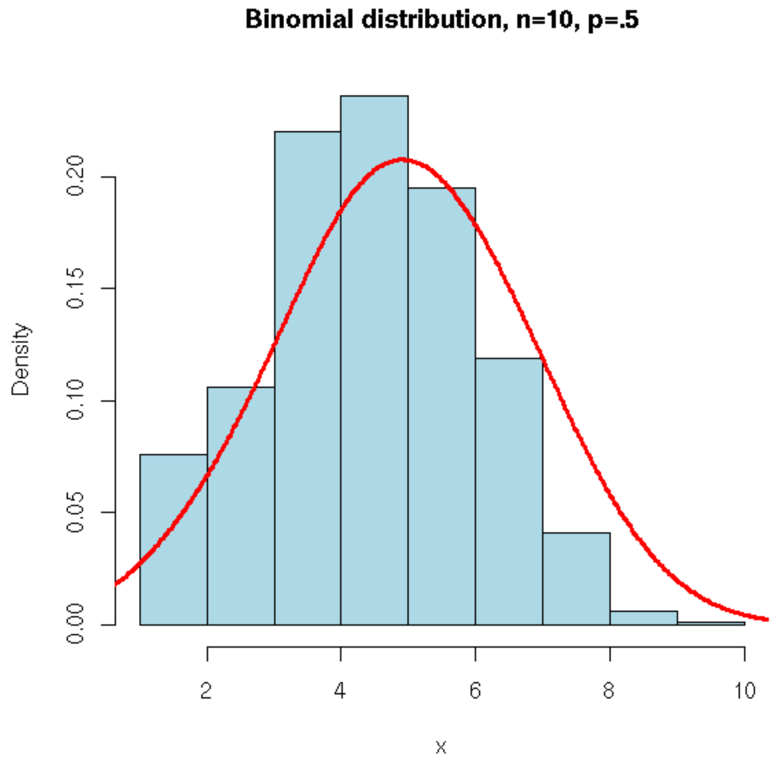


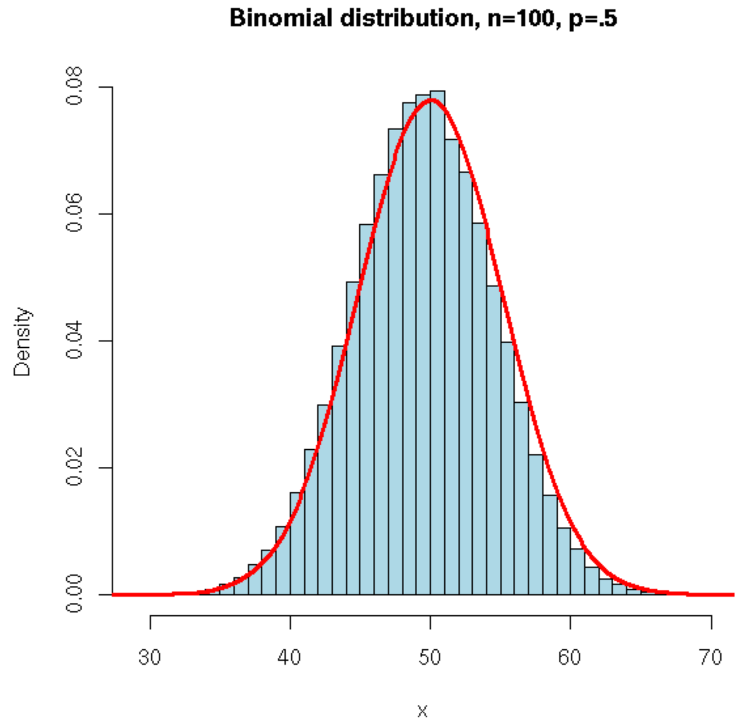
$$f(k;r,p) = \frac{\Gamma(r+k)}{k!\,\Gamma(r)} p^r (1-p)^k$$

- Characteristics of Bernoulli exp.
  - Repeated trials
  - Each trial has two possible outcomes: success/failure
  - Probability that success occurs is same on every trial
  - Trials are independent

Binomial distribution

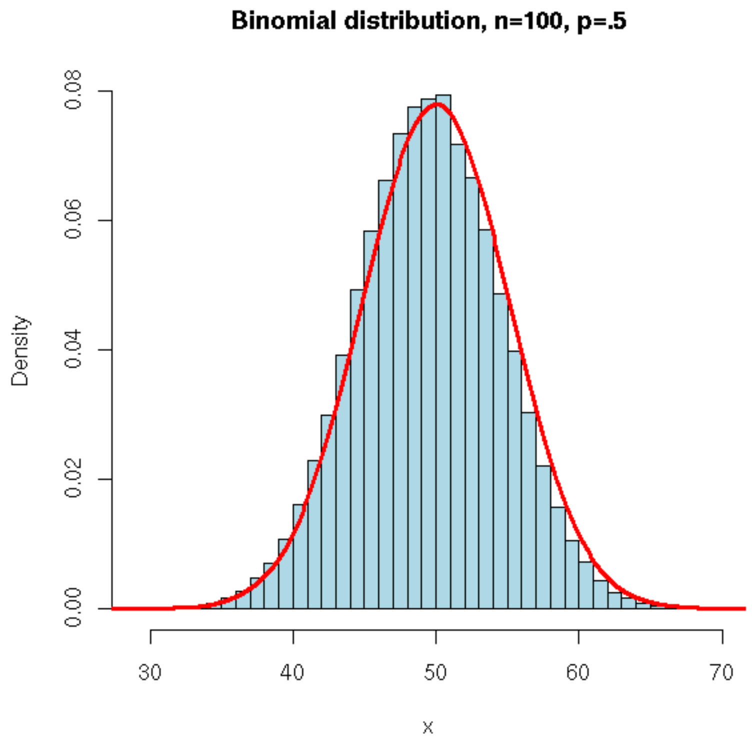
- Suppose we toss a coin n times and count the number of “heads”.
- Probability is constant – 0.5 on each trial
- B(x; n,p) is the binomial probability: the probability that n-trials results in exactly x successes when the probability in each trial is p

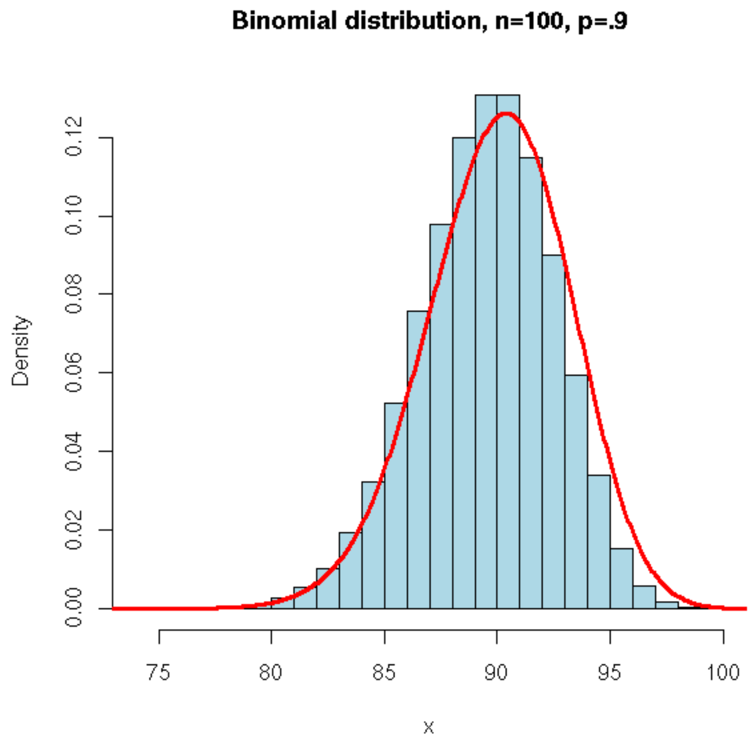




Binomial distribution

- Properties of the binomial distribution:
  - Mean of the distribution is $\mu_x = n*p$
  - Variance is $\sigma^2_x = n*p*(1-p)$
  - Standard deviation is: $\sigma_x = \mathrm{sqrt}[n*p*(1-p)]$

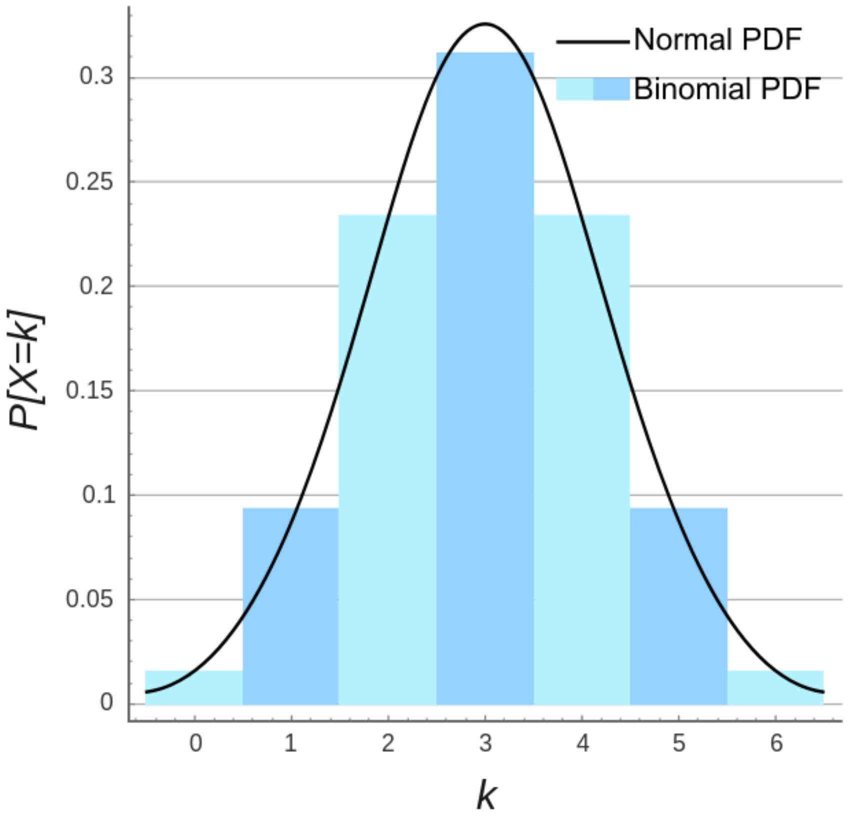




Binomial distribution

- If n is large enough then a good approximation to B(n,p) is the *normal distribution*

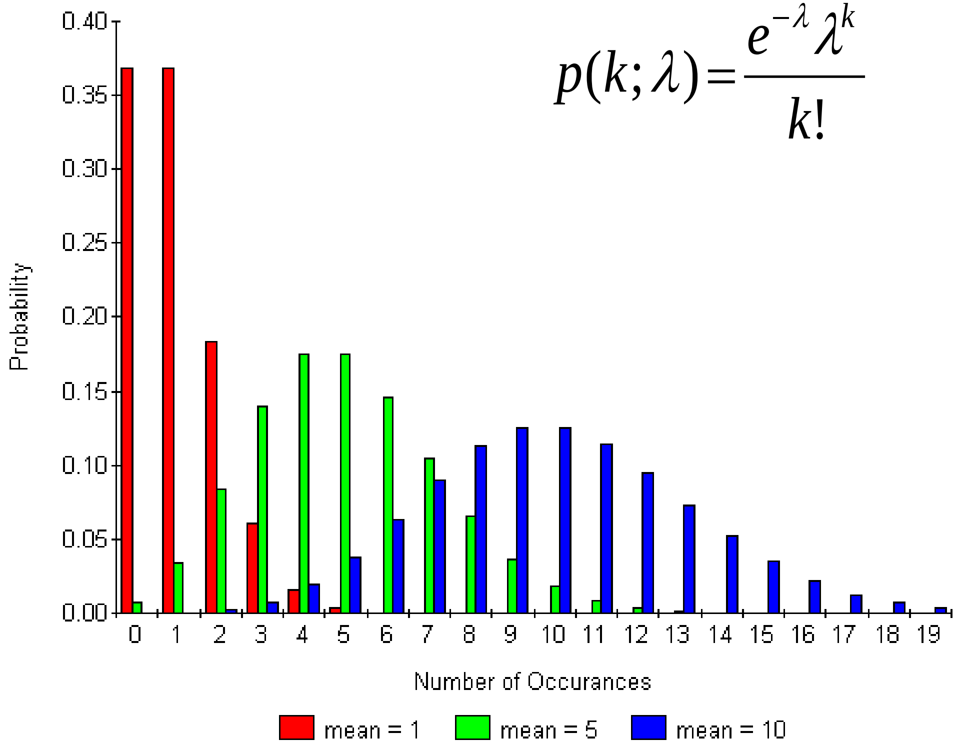
- Approximation improves as n increases and is better if p is not near 0 or 1



Binomial and normal approximation for n=6 and p=0.5

Poisson distribution

- Binomial distribution converges towards the Poisson distribution as the number of trials goes to infinity while the product *np* remains fixed.
  - Poisson distribution with parameter $\lambda = np$ can be used as an approximation to B(n, p) of the binomial distribution if n is sufficiently large and p is sufficiently small, i.e., if $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$



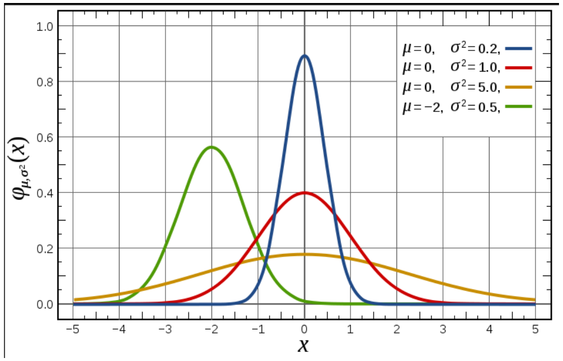
$$p(k;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$

k events with an expected number of events λ

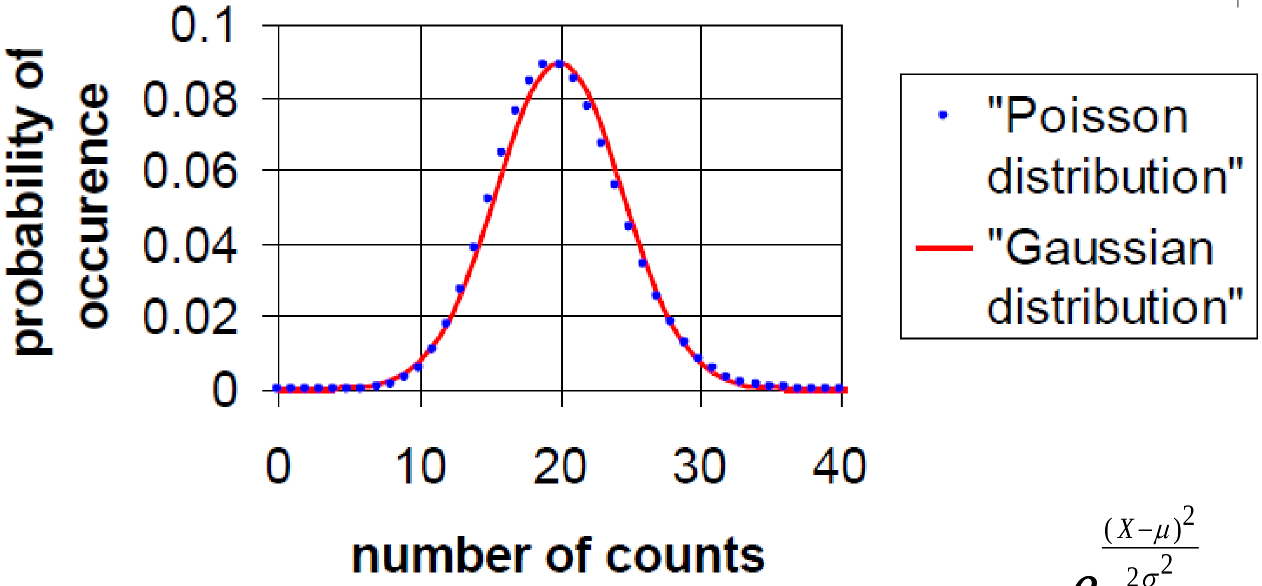
- Characteristics of Poisson exp.
  - Outcomes are success/failure
  - Average number of successes in specified region is known
  - Probability that success occurs is proportional to size of region
  - Probability that success occurs in very small region is zero

Poisson distribution at large λ

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!}$$ k=0,1,2,...



Poisson and Gaussian distributions



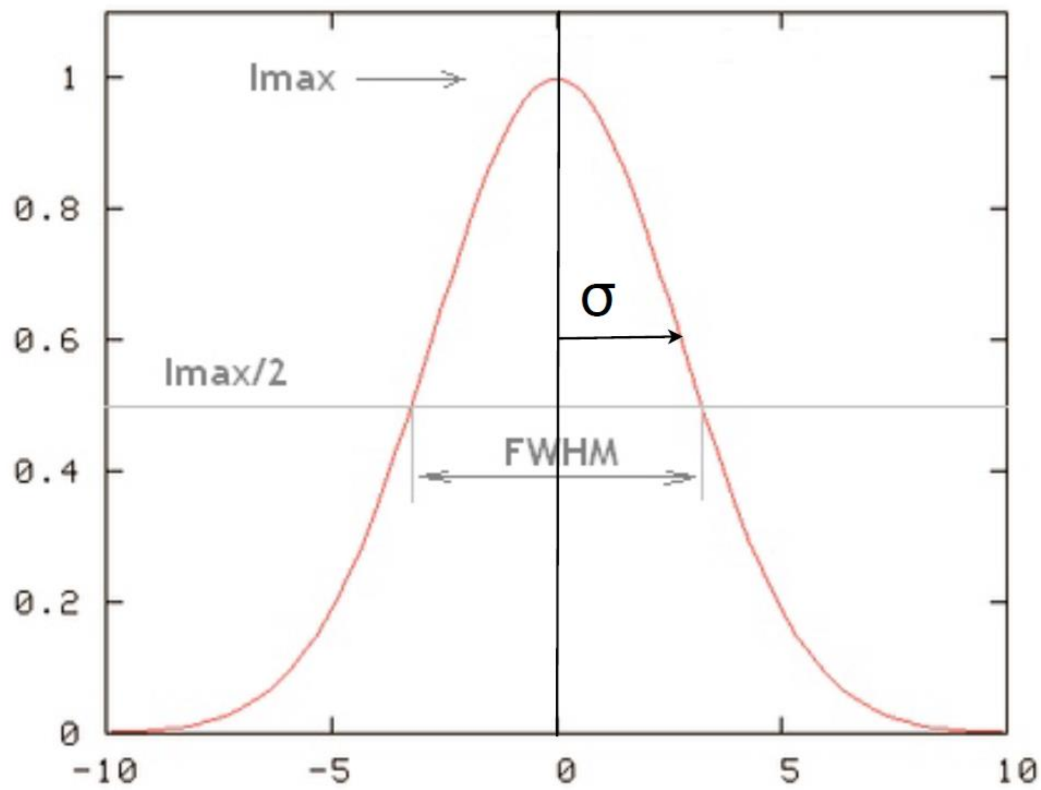
$$p_k = \frac{e^{\frac{(X-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

Gaussian distribution: continous

Gaussian distribution

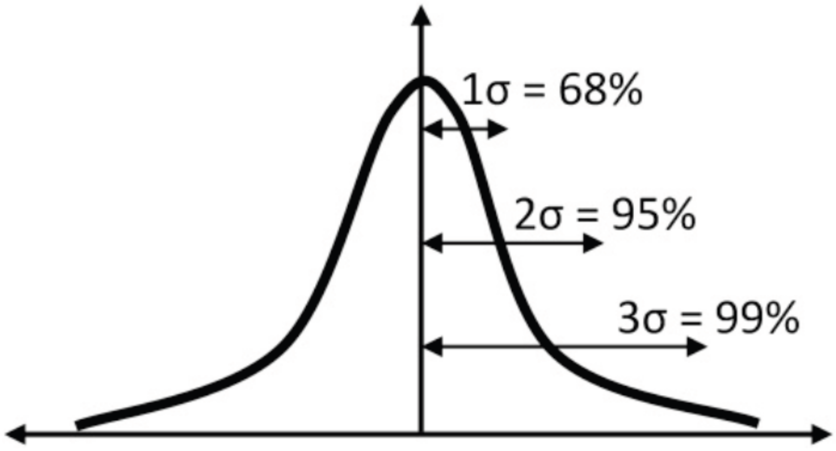
- Graph depends on:
  - Mean, μ
  - Standard deviation, σ

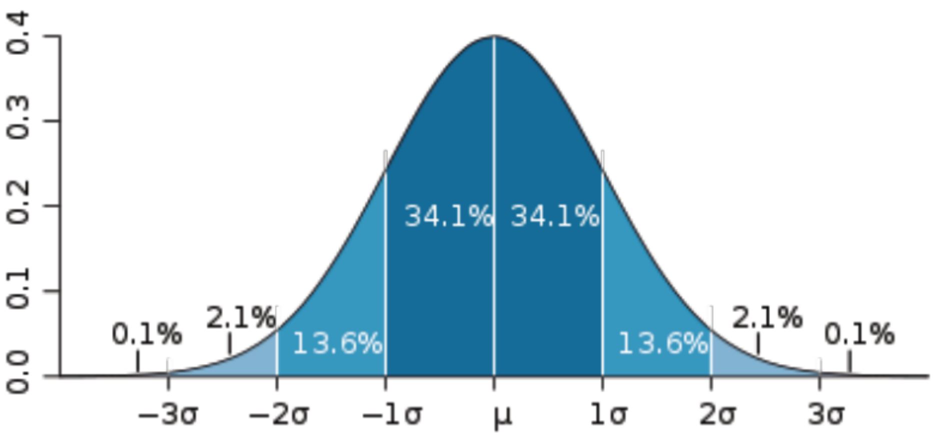
$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Gaussian distribution

- For every normal curve:
  - About 68% of the area under the curve falls within 1 standard deviation of the mean
  - About 95% of the area under the curve falls within 2 standard deviation of the mean
  - About 99% of the area under the curve falls within 3 standard deviation of the mean





Chi2 distribution

- Distribution of a sum of squares of k independent standard normal random variables
- Contruction of confidence intervals,
- Test goodness of fit of an observed distribution to a theoretical one