TRIUMF

Detector Physics

Statistical Methods

November 30, 2015

D. Lascar | Postdoctoral Fellow | TRIUMF

Outline

- Probability Distributions
- Uncertainty Measurements and Propagation
- Curve Fitting

Cumulative distribution

- Lets say you have a probability distribution

$$P(x')$$

Probability of finding a thing at x'

Cumulative distribution

Lets say you have a probability distribution

$$P(x')$$

$$P(x)dx$$

Probability of finding x' in the interval $x + dx$

Cumulative distribution

Lets say you have a probability distribution

$$P(x')$$

$$P(x)dx$$

Normally you want to find x' within certain limits

$$P(x_1 \leq x' \leq x_2) = \begin{cases} \int_{x_1}^{x_2} P(x)dx \\ \sum_{i=1}^{2} P(x_i) \end{cases}$$

Cumulative distribution

Lets say you have a probability distribution

$$P(x')$$

$$P(x)dx$$

Normally you want to find x' within certain limits

$$P(x_1 \leq x' \leq x_2) = \begin{cases} \int_{x_1}^{x_2} P(x)dx \rightarrow \int P(x)dx = 1 \\ \sum_{i=1}^{2} P(x_i) \rightarrow \sum_{i} P(x_i) = 1 \end{cases}$$

Expectation values

If x' is a random variable distributed as $P(x)$

$$E[x'] = \langle x' \rangle = \begin{cases} \int x' P(x) dx \\ \sum_i x' P(x_i) \end{cases}$$

$$E[f(x')] = \langle f(x') \rangle = \int f(x') P(x) dx$$

Expectation values

If x' is a random variable distributed as $P(x)$

$$E[x'] = \langle x' \rangle = \begin{cases} \displaystyle\int x' P(x) dx \\ \displaystyle\sum_i x' P(x_i) \end{cases}$$

$$E[f(x')] = \langle f(x') \rangle = \int f(x') P(x) dx$$

This leads us to the theoretical mean

$$\bar{x} = \int x P(x) dx$$

Expectation values

The variance

$$\sigma^2 = \langle (x - \bar{x})^2 \rangle = \int (x - \bar{x})^2 P(x) dx$$

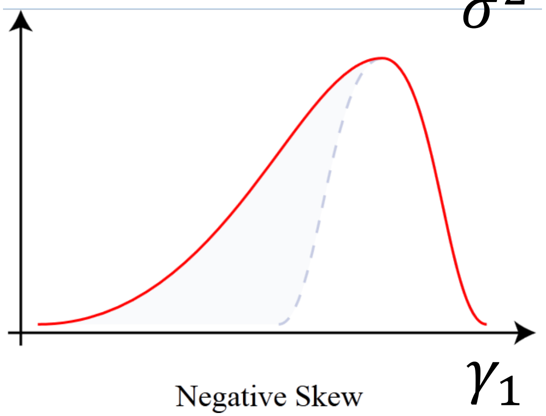
This leads us to the theoretical mean

$$\bar{x} = \int x P(x) dx$$

Expectation values

The variance

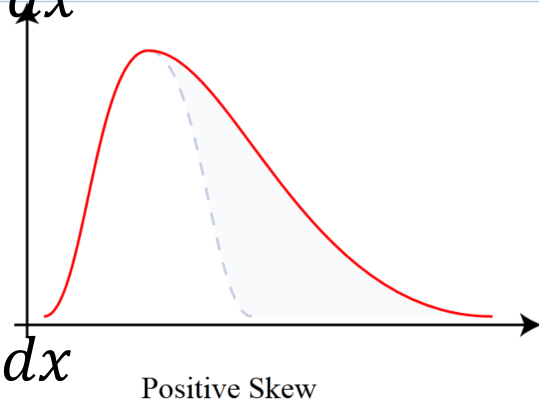
$$\sigma^2 = \langle (x-\bar{x})^2 \rangle = \int (x-\bar{x})^2 P(x)dx$$



Source: Wikimedia commons

Skewness

$$\gamma_1 = \langle (x-\bar{x})^3 \rangle = \int (x-\bar{x})^2 P(x)dx$$



Source: Wikimedia commons

This leads us to the theoretical mean

$$\bar{x} = \int xP(x)dx$$

Adding dimensions/variables

- So far we've only dealt with 1-D probability
  - Life is rarely that simple
- $P(x, y, z, \cdots)$ is much more likely
- Everything is defined as before but integrated over all directions/variables

$$\langle x' \rangle = \int x' P(x, y, z, \cdots) dx dy dz \cdots$$

$$\sigma^2(x) = \langle (x - \bar{x})^2 \rangle = \int (x - \bar{x})^2 P(x, y, z, \cdots) dx dy dz \cdots$$

$$\gamma_1(x) = \langle (x - \bar{x})^3 \rangle = \int (x - \bar{x})^3 P(x, y, z, \cdots) dx dy dz \cdots$$

Covariance

- Measure of linear correlation between 2 variables $\mathrm{cov}(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle$

Covariance

- Measure of linear correlation between 2 variables $\mathrm{cov}(x,y) = \langle (x-\bar{x})(y-\bar{y}) \rangle$
- If $P(x,y,z)$
  - $\mathrm{cov}(x,y)$
  - $\mathrm{cov}(x,z)$
  - $\mathrm{cov}(y,z)$
- Normally expressed as a correlation coefficient
  - $|\rho| = 1 \rightarrow$ correlated linearly
  - $\rho = 0 \rightarrow$ variables are linearly independent

$$\rho = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$$

↓

$$P(x,y) = P(x)P(y)$$

Binomial distribution

- You have two possible outcomes
  - Heads v. Tails, Yes v. No, Hit v. Miss
  - Probability the unchanged between trials
  - Not necessarily 50-50

$$P(r) = \frac{N!}{r!\,(N-r)!}\,p^r(1-p)^{N-r}$$

Number of successes

Number of trials

Probability of success from an individual trial

Binomial distribution

- You have two possible outcomes
  - Heads v. Tails, Yes v. No, Hit v. Miss
  - Probability the unchanged between trials
  - Not necessarily 50-50

$$P(r) = \frac{N!}{r!\,(N-r)!} p^r (1-p)^{N-r}$$

- This is (by definition) a discrete distribution so

$$\bar{r} = \sum rP(r) = Np$$

$$\sigma^2 = \sum (r-\bar{r})^2 P(r) = Np(1-p)$$

Normalizing the binomial distribution

The goal is to make $\sum P(r) = 1$ so

$$\sum_{r=0}^{N} P(r) = \sum_{r=0}^{N} \frac{N!}{r!\,(N-r)!} p^r (1-p)^{N-r}$$

*r*th term of the binomial expansion

Normalizing the binomial distribution

The goal is to make $\sum P(r) = 1$ so

$$\sum_{r=0}^{N} P(r) = \sum_{r=0}^{N} \frac{N!}{r!\,(N-r)!} p^r (1-p)^{N-r} = [(1-p)+p]^N = 1$$

Normalizing the binomial distribution

The goal is to make $\sum P(r) = 1$ so

$$\sum_{r=0}^{N} P(r) = \sum_{r=0}^{N} \frac{N!}{r!\,(N-r)!} p^r (1-p)^{N-r} = [(1-p)+p]^N = 1$$

For many trials ($N \rightarrow$ big)

and

$p \leq 0.05$ & Np finite

Use the Poisson Distribution

The Poisson distribution

- Limiting form of the binomial distribution
  - $N \to \infty$
  - $p \to 0$ $\qquad \bar{r} = Np$ is finite
  - $0 < \bar{r} \ll \infty$

The Poisson distribution

- Limiting form of the binomial distribution
  - $N \to \infty$
  - $p \to 0$ $\quad\quad\bar{r} = Np$ is finite
  - $0 < \bar{r} \ll \infty$
- Probability of observing *r* events

$$P(r) = \frac{\bar{r}^r e^{-\bar{r}}}{r!}$$

- This is appropriate for
  - Nuclear reactions
  - Radioactive decay

Example: 1 μg ^{137}Cs

- $t_{1/2}$ = 27 years
  - $\lambda = \frac{\ln 2}{27} = 0.026\,\text{years} = 8.2 \times 10^{-10}\ \text{s}^{-1}$
- $1\ \mu g = 4.4 \times 10^{15}$ Cs atoms
- $N\lambda = 3.6 \times 10^{6}$ decays/s

p is small

N is large

Np is finite

Example: 1 μg ^{137}Cs

- $t_{1/2}$ = 27 years
  - $\lambda = \frac{\ln 2}{27} = 0.026 \text{ years} = 8.2 \times 10^{-10}\ \text{s}^{-1}$
- $1\ \mu\text{g} = 4.4 \times 10^{15}$ Cs atoms
- $N\lambda = 3.6 \times 10^{6}$ decays/s
- We can use:

$$P(r) = \frac{\bar{r}^{r} e^{-\bar{r}}}{r!}$$

- But only if we realize that $\bar{r} = \lambda t$

p is small

N is large

Np is finite

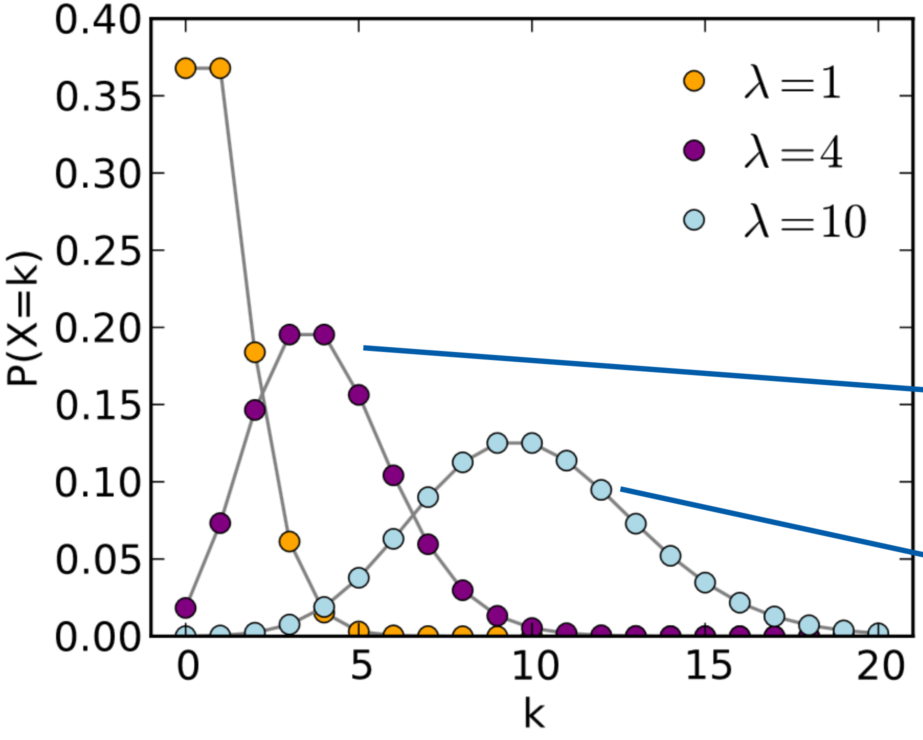
Decays and reactions

$$\bar{r} = \lambda t$$

$$P(r) = \frac{\bar{r}^r e^{-\bar{r}}}{r!} = \frac{(\lambda t)^r e^{-\lambda t}}{r!}$$

$$\sigma^2 = \lambda t$$

Prove for homework



The distribution changes as a function of time

Not symmetric

Getting more symmetric

Normalizing the binomial distribution

The goal is to make $\sum P(r) = 1$ so

$$\sum_{r=0}^{N} P(r) = \sum_{r=0}^{N} \frac{N!}{r!\,(N-r)!} p^r (1-p)^{N-r} = [(1-p)+p]^N = 1$$

For many trials ($N \rightarrow$ big)

and

$p \leq 0.05$ & Np finite

Use the Poisson Distribution

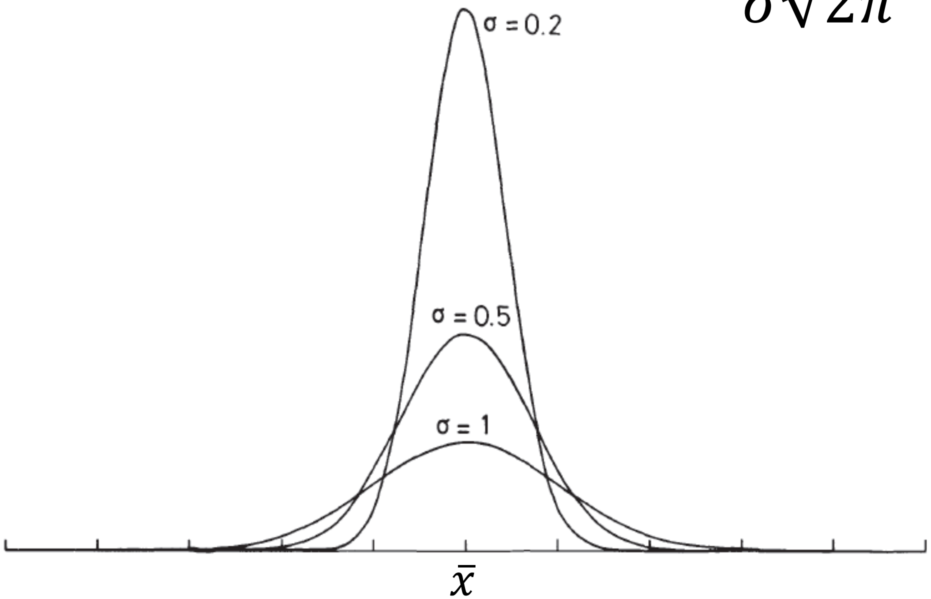
$p \geq 0.05$ & Np finite

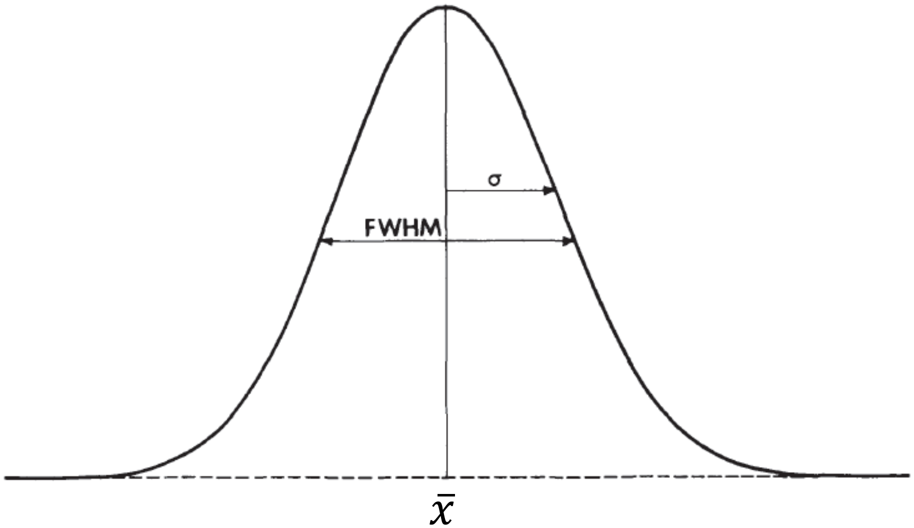
Use the Gaussian Distribution

The Gaussian (normal) distribution

- The most ubiquitous distribution in all the sciences.
- Even when it's application isn't the best it is still used.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$



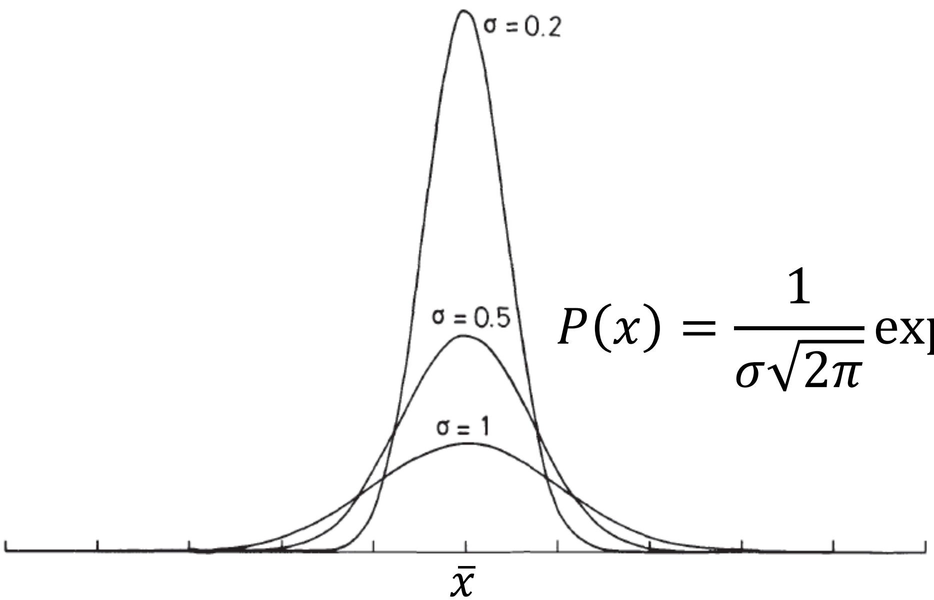


The Gaussian (normal) distribution

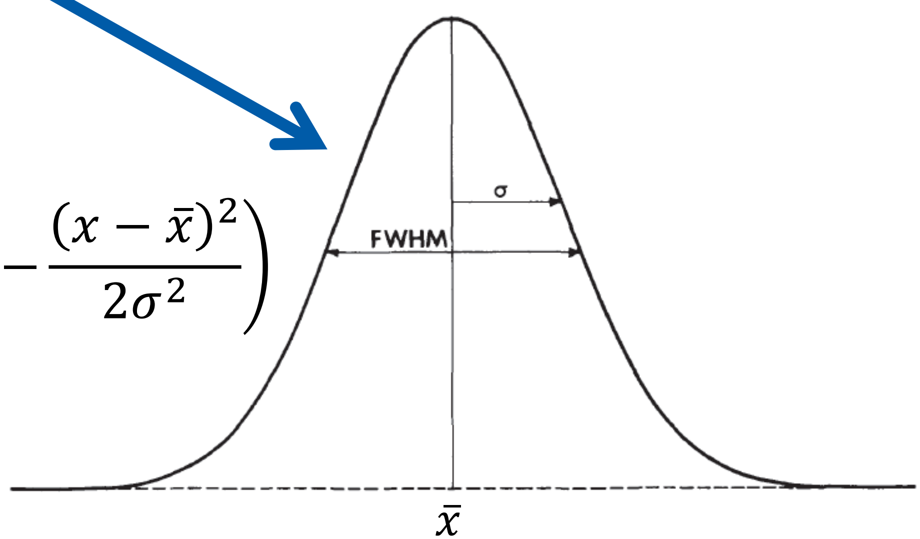
- FWHM isn't σ

$$\mathrm{FWHM} = 2\sigma\sqrt{\ln 4} = 2.35\sigma$$

Prove for homework



$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$



The Gaussian (normal) distribution

- FWHM isn't σ

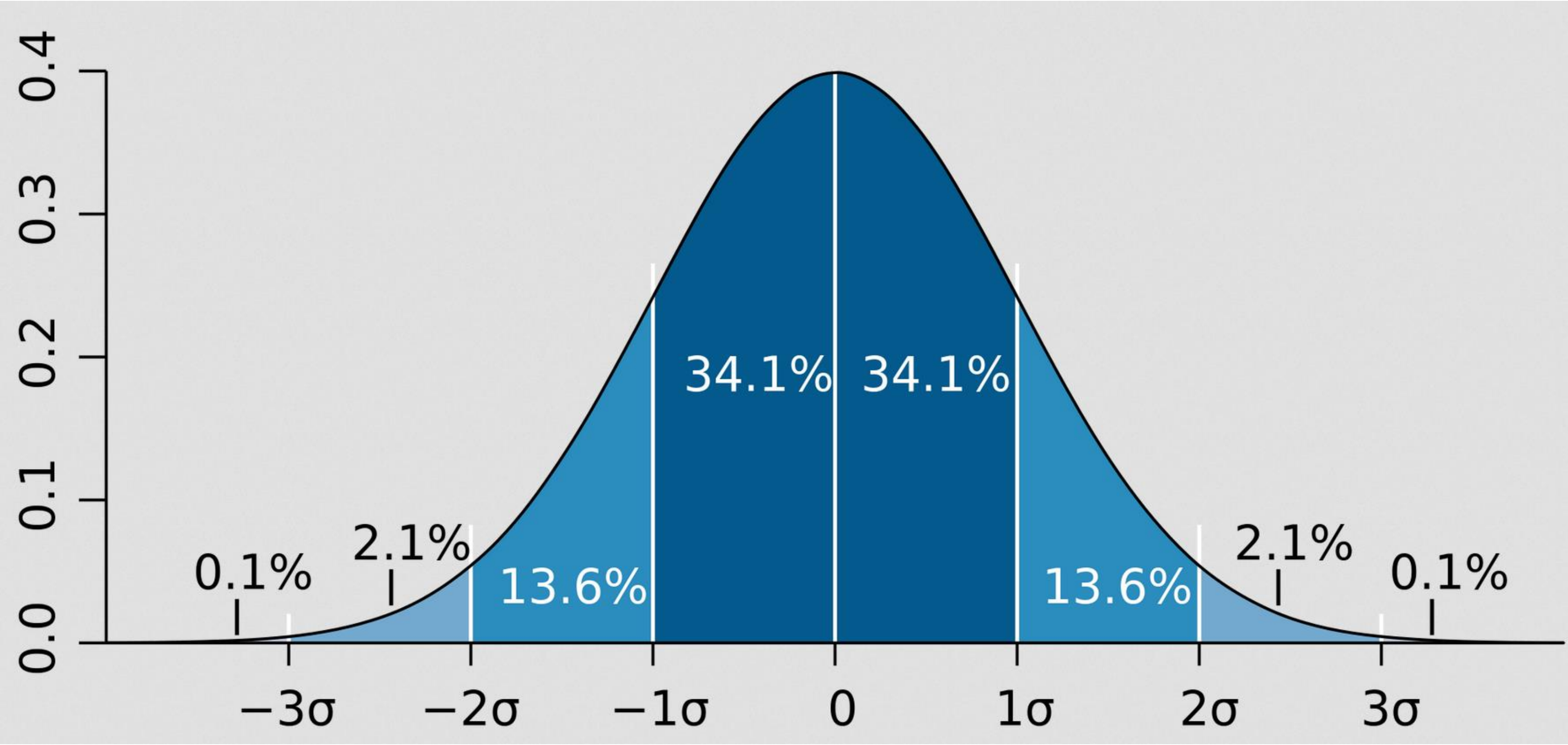
$$\mathrm{FWHM} = 2\sigma\sqrt{\ln 4} = 2.35\sigma$$

- No analytical solution → Numerical methods
- Tables tend to be calculated for reduced Gaussian
  - $\bar{x} = 0$
  - $\sigma^2 = 1$

$$\left.\begin{array}{l} \bar{x} = 0 \\ \sigma^2 = 1 \end{array}\right\} z = \frac{x - \bar{x}}{\sigma}$$

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

Standard deviation (σ)



Uncertainty Measurements

Systematic Uncertainty

- Come from a bias in the data
- Tend to be in one direction
- Difficult to account
- Different for every experimental setup
- Should be minimized whenever possible/practical

Random Uncertainty

- Result from:
  - Statistical fluctuations in the data
  - Random imprecisions in the measurement device
- Are random in direction
- Determined via sampling
- Should be minimized whenever possible/practical

Sampling

- Given a sample of measurements
  - $x_1, x_2, x_3, \ldots, x_n$

$$\overline{x'} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Sampling estimation

- Given a sample of measurements
  - $x_1, x_2, x_3, \dots, x_n$

$$\overline{x'} = \frac{1}{n}\sum_{i=1}^{n} x_i \neq \int xP(x)dx = \bar{x}$$

$$\overline{x'} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} x_i = \int xP(x)dx = \bar{x}$$

- Similarly $\lim_{n\to\infty} s^2 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x'})^2 = \sigma^2$

- Larger sample sizes approach the theoretical value

Sampling estimation – Maximum likelihood method

- Only possible if the probability distribution of the sample is known
- Given:
  - n independent observations $(x_1, x_2, x_3, \ldots, x_n)$
  - Probability distribution $f(x|\theta)$
- The goal is to calculate a *Likelihood function*

$$L(\theta|x) = f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta)$$

Probability of observing the sequence $x_1, x_2, x_3, \ldots, x_n$

Sampling estimation – Maximum likelihood method

- Only possible if the probability distribution of the sample is known
- Given:
  - n independent observations $(x_1, x_2, x_3, \ldots, x_n)$
  - Probability distribution $f(x|\theta)$
- The goal is to calculate a *Likelihood function*

$$L(\theta|x) = f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta)$$

- $L(\theta|x)$ defined to be a maximum for $(x_1, \ldots, x_n)$

$$\frac{dL}{d\theta} = 0$$

Sampling estimation – Maximum likelihood method

- $L(\theta|x)$ defined to be a maximum for $(x_1, \dots, x_n)$

$$\frac{dL}{d\theta} = 0$$

- Solution: $\hat{\theta}$ – *Maximum likelihood estimator*

$$\sigma^2(\hat{\theta}) = \int (\hat{\theta} - \theta)^2 L(\theta|x)dx_1 \cdots dx_n$$

- Only solvable (analytically) in a few simple cases
- For solutions with Poisson distributions
  - See Leo 4.4.3
- For solutions with Gaussian distributions
  - See Leo 4.4.4

Propagation of uncertainties

- Consider $u = f(x, y)$ with σ_x & σ_y
- The goal is to calculate $\sigma_u(\sigma_x, \sigma_y)$

$$\sigma_u^2 = \langle (u - \bar{u})^2 \rangle$$

Propagation of uncertainties

- Consider $u = f(x, y)$ with σ_x & σ_y
- The goal is to calculate $\sigma_u(\sigma_x, \sigma_y)$

$$\sigma_u^2 = \langle (u - \bar{u})^2 \rangle$$

- To 1st order, $\bar{u} \approx f(\bar{x}, \bar{y})$

$$u - \bar{u} \cong (x - \bar{x}) \left.\frac{\partial f}{\partial x}\right|_{\bar{x}} + (y - \bar{y}) \left.\frac{\partial f}{\partial y}\right|_{\bar{y}}$$

$$\langle (u - \bar{u})^2 \rangle \cong$$

$$\left\langle (x - \bar{x})^2 \left(\frac{\partial f}{\partial x}\right)^2 + (y - \bar{y})^2 \left(\frac{\partial f}{\partial y}\right)^2 + 2(x - \bar{x})(y - \bar{y}) \frac{\partial^2 f}{\partial x \partial y} \right\rangle$$

Propagation of uncertainties

$$\sigma_u^2 = \langle (u - \bar{u})^2 \rangle \cong$$

$$\cong \left\langle (x - \bar{x})^2 \left(\frac{\partial f}{\partial x}\right)^2 + (y - \bar{y})^2 \left(\frac{\partial f}{\partial y}\right)^2 + 2(x - \bar{x})(y - \bar{y}) \frac{\partial^2 f}{\partial x \partial y} \right\rangle$$

- Take the expectation value of each term separately:

$$\sigma_u^2 \cong \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\mathrm{cov}(x, y) \frac{\partial^2 f}{\partial x \partial y}$$

Propagation of uncertainties - cases

- Sums (and differences)
  - $q = x + \cdots + z - (u + \cdots + w)$

$$\sigma_q \begin{cases} = \sqrt{\sigma_x^2 + \cdots + \sigma_z^2 + \sigma_u^2 + \cdots + \sigma_w^2} \\ \leq \quad \sigma_x + \cdots + \sigma_z + \sigma_u + \cdots + \sigma_w \end{cases}$$

- Products (and quotients)
  - $q = \frac{x \times \cdots \times z}{u \times \cdots \times w}$

$$\frac{\sigma_q}{|q|} \begin{cases} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \cdots + \left(\frac{\sigma_z}{z}\right)^2 + \left(\frac{\sigma_u}{u}\right)^2 + \cdots + \left(\frac{\sigma_w}{w}\right)^2} \\ \leq \quad \frac{\sigma_x}{|x|} + \cdots + \frac{\sigma_z}{|z|} + \frac{\sigma_u}{|u|} + \cdots + \frac{\sigma_w}{|w|} \end{cases}$$

Propagation of uncertainties - cases

- If $q = Bx$ and B is a known constant
  - $\sigma_q = |B|\sigma_x$
- If $q = x^n$
  - $\frac{\sigma_q}{|q|} = \frac{|n|\partial x}{|x|}$
- Proofs of all of these are in:
  - J.R. Taylor, *An Introduction to Error Analysis: The study of uncertainties in physical measurements*," 2nd Ed. Sausalito, Ca, University Science Books 1997.

Curve fitting

- Most of the time we measure some value as a function of several variables that we set

$$u_i = f(x_i, y_i, \cdots)$$

- We need to fit these points to a theoretical curve that describes the behavior
- Example: Radioactive source
  - Measure count rates: $N_1, N_2, \ldots, N_n$
  - Measurements made at times: $t_1, t_2, \ldots, t_n$
- Data should fit: $N(t) = N_0 \exp({}^{-t}/_{\tau})$
- What is the best way to find N_0 & τ?

Least squares fitting

- Measure n points at x_i of y_i with error σ_i
- Need to fit $f(x; a_1, a_2, \ldots, a_m)$ which should describe y
  - $a_1, a_2, \ldots, a_m$ are unknown parameters to be solved

$$\boldsymbol{n > m}$$

- Best values of a_j:

$$S = \sum_{i=1}^{n} \left[\frac{y_i - f(x_i | a_1, \ldots, a_m)}{\sigma_i} \right]^2$$

Minimize

Chi-squared (χ^2) distribution

Least squares fitting

- Measure n points at x_i of y_i with error σ_i
- Need to fit $f(x; a_1, a_2, \ldots, a_m)$ which should describe y
  - $a_1, a_2, \ldots, a_m$ are unknown parameters to be solved

$$\boldsymbol{n > m}$$

- Best values of a_j:

$$S = \sum_{i=1}^{n} \left[\frac{y_i - f(x_i | a_1, \ldots, a_m)}{\sigma_i} \right]^2$$

A system of m equations

$$\frac{\partial S}{\partial a_j} = 0$$

Least s[redacted]es fitting

$$\frac{\partial S}{\partial a_j} = 0$$

- Life is [redacted] rarely that easy.
- If there's no analytic solution (most of the time) we need numerical methods to minimize S
- We create the covariance or *error matrix*

$$\left(\tilde{V}^{-1}\right)_{kl} = \frac{1}{2}\frac{\partial^2 S}{\partial a_k \partial a_l}$$

Minimize

Least s[redacted]es fitting

$$\frac{\partial S}{\partial a_j} = 0$$

- Life is [redacted] rarely that easy.
- If there's no analytic solution (most of the time) we need numerical methods to minimize S
- We create the covariance or *error matrix*

$$\left(\tilde{V}^{-1}\right)_{kl} = \frac{1}{2}\frac{\partial^2 S}{\partial a_k \partial a_l}$$

Least s[redacted]es fitting

$$\frac{\partial S}{\partial a_j} = 0$$

- Life is [redacted] rarely that easy.
- If there's no analytic solution (most of the time) we need numerical methods to minimize S
- We create the covariance or *error matrix*

$$\left(\tilde{V}^{-1}\right)_{kl} = \frac{1}{2}\frac{\partial^2 S}{\partial a_k \partial a_l}$$

$$\tilde{V} = \begin{bmatrix} \sigma_1^2 & \mathrm{cov}(a_1, a_2) & \mathrm{cov}(a_1, a_3) & \cdots \\ \cdot & \sigma_2^2 & \mathrm{cov}(a_2, a_3) & \cdots \\ \cdot & \cdot & \sigma_3^2 & \cdots \\ \cdot & \cdot & \cdot & \ddots \end{bmatrix}$$

Nonlinear fits

- Sadly nonlinear functions are too difficult for this class length.
- References:
  - W.T. Eadie, *et al*: *Statistical Methods in Experimental Physics* (North-Holland, Amsterdam, London 1971)
    - QC39 .S74 1971 – In UBC library
  - F. James: *Statistical Methods in Experimental Physics* (World Scientific, Hackensack 2006)
    - QC39 .S74 2006 – In TRIUMF library
  - P.R. Bevington, *et* al: *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, Boston 2003)
    - QA278 .B48 2003 – In UBC Library

Thank you!
Merci

TRIUMF

Canada's national laboratory for particle and nuclear physics

Laboratoire national canadien pour la recherche en physique nucléaire et en physique des particules

TRIUMF: Alberta | British Columbia | Calgary | Carleton | Guelph | Manitoba | McGill | McMaster | Montréal | Northern British Columbia | Queen's | Regina | Saint Mary's | Simon Fraser | Toronto | Victoria | Western | Winnipeg | York