PHYS 575

Autumn 2015

Radiation and Radiation Detectors

Course home page:

http://depts.washington.edu/physcert/radcert/575website/

5: cross sections, attenuation; calorimetry; counting statistics; statistics for analysis

R. Jeffrey Wilkes

Department of Physics

B305 Physics-Astronomy Building

206-543-4232

wilkes@u.washington.edu

Course calendar

week	date	day	topic	text
1	10/1/15	Thurs	Introduction, review of basics, radioactivity, units for radiation and dosimetry	Ch. 1, notes
2	10/6/15	Tues	Radioactive sources; decay processes;	Ch. 1, notes
3	10/13/15	Tues	Photomultiplier tubes and scintillation counters; Counting statistics	Chs. 3, 8, 9 (I-V)
3	10/15/15	Thurs	**LAB: Room B248** Scopes, fast pulses; PMTs and scintillation counters; standard electronics modules	Chs. 4, 9, 16, 17
4	10/20/15	Tues	Overview of charged particle detectors	Ch. 4
4	10/22/15	Thurs	**LAB: Room B248** Coincidence techniques; nanosec time measurement, energy from pulse area	Chs. 17, 18
5	10/27/15	Tues	Interaction of charged particles and photons with matter; counting statistics; gas detectors; Proposal for term paper must be emailed to JW by today	Chs. 2, 3; Chs. 5, 6, 7
6	11/3/15	Tues	ionization chambers; solid-state detectors	Chs. 11, 12, 13
7	11/10/15	Tues	Detecting neutral particles; Data acquisition methods	Ch. 14, 15, 18
8	11/17/15	Tues	Cherenkov detectors; Case studies: neutrino detectors (IceCube), atmospheric Cherenkov, triggering Cherenkov	Ch. 19, notes
9	11/24/15	Tues	Case studies: classic detectors (cloud and bubble chambers, nuclear emulsion), high energy accelerators, Fermi LAT	Ch. 19, notes
10	12/1/15	Tues	Finish case studies; begin student presentations	Notes
11	12/8/15	Tues	Student presentations	-
11	12/10/15	Thurs	Student presentations	

Tonight (→ week 5)

Announcements

- Send me your report proposal! topic, brief summary, list of proposed resources/references
- Presentation dates: Tues Dec 1, Tues Dec 8, and Thurs Dec 10
  - See class web page for link to signup sheet

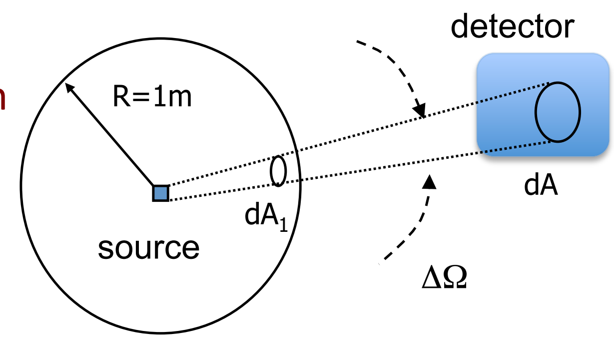
- NEW Schedule and signup table for term project presentations. This is a Google spreadsheet in the UW Google Docs filespace; log in with your UW NetID username and password (NOT your personal Google username) for access. Sign in to the slot you want, then exit, and let me know you did so by email.

  - Volunteers needed for early slots (Dec 1)
  - I will arbitrarily assign slots for those not signed up by November 29

Cross sections and accceptance

- First: Ω = solid angle acceptance of detector:

  Project an element of detector area as viewed from source onto a 1m radius sphere:



- Total cross section $\sigma = N_{events} / (\Delta t \, N_{tgt} \, I_{beam})$

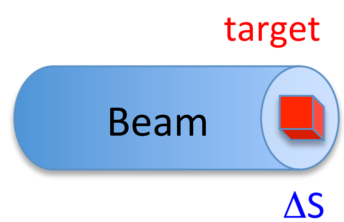
  $I_{beam} = (N_{beam} / \Delta t) / \Delta S$ for beam area ΔS larger than target, particles/cm^2/sec

  N_{tgt} = number of target nuclei in beam area

  (or number of electrons in beam, for Compton scattering)

  $= (N_{avogodro} \, M_{tgt} / A_{tgt}) Z_{tgt}$ (M_{tgt} in moles)

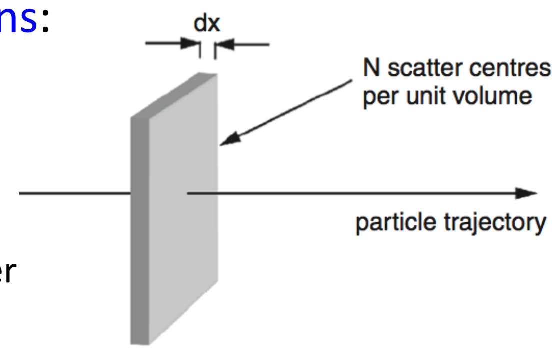
  σ has dimensions of area, unit : barn = 10^{-28} m^2

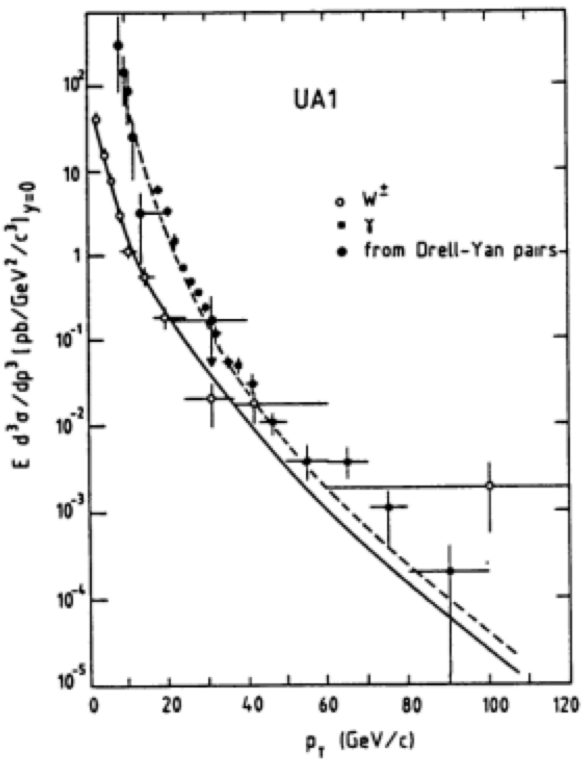


- Can have various partial differential cross sections:
  - vs energy = $d\sigma/dE$
  - vs angle = $d\sigma/d\theta$
  - vs several things at once: eg, $d^3\sigma / dE \, d\theta \, d\Omega$

    = cross section at energy E (per unit E) and angle θ (per unit θ), per unit solid angle $d\Omega$ centered on θ
  - area under $d\sigma/d$(whatever) curve = σ

    total cross section for that process





Partial differential cross sections
From 1988 Int. Conf. on HEP
(UA1 experiment reports)

Fig. 2 The invariant cross section for W and direct photon production as a function of transverse momentum. The theoretical curves are from Altarelli et al. [7] and Aurenche et al. [13].

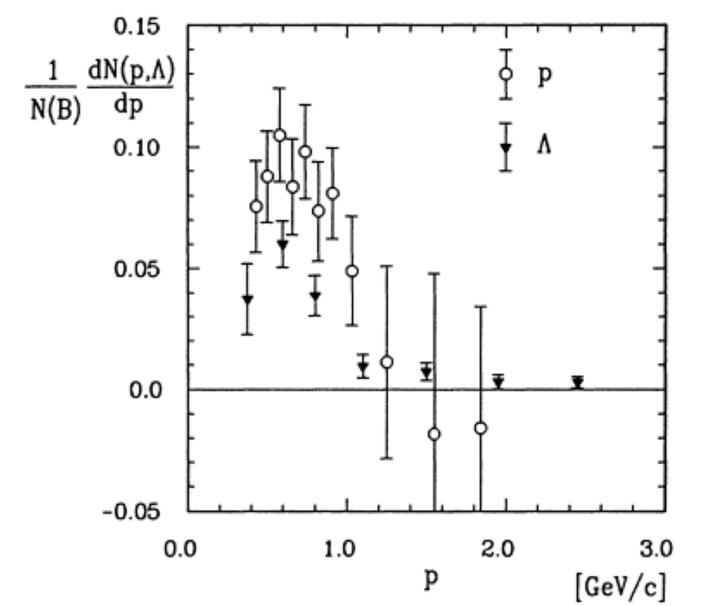


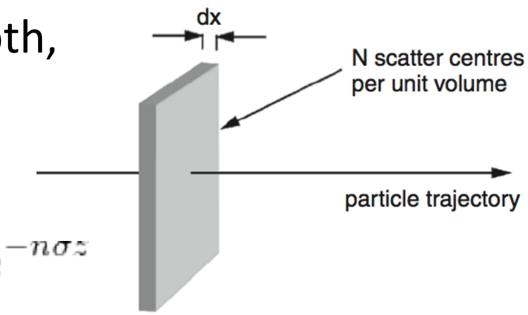
Figure 3: Momentum spectrum of p and Λ from B meson decays.

Attenuation length

- Attenuation length:

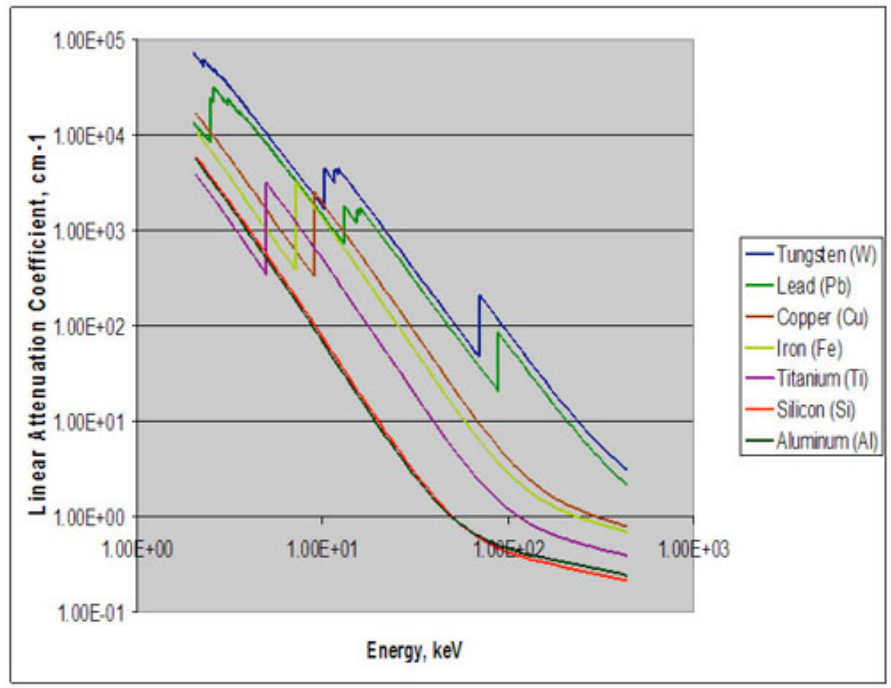
  In terms of flux Φ (particles per cm^2 in beam) vs depth,

$$\frac{\mathrm{d}\Phi}{\mathrm{d}z} = -n\sigma\Phi$$



  Simple differential eqn. which has solution $\Phi = \Phi_0 \mathrm{e}^{-n\sigma z}$

  n = number density of targets in 1/cm^{-3}
  z = thickness in g/cm^2
  ns = 1/λ, attenuation length in cm



- Attenuation coefficient:

$$\mu = -\frac{1}{\Phi_\mathrm{e}}\frac{\mathrm{d}\Phi_\mathrm{e}}{\mathrm{d}z},$$

  – So $\mu = n\sigma = 1/\lambda$

  Mass attenuation coefficient:

$$\frac{\mu}{\rho_m},$$

Collisions: impact parameter

- Impact parameter b

  Distance between target center and beam particle in individual collision

  – Example: Rutherford scattering: he found $\frac{d\sigma}{d\Omega} = \frac{b}{\sin\Theta}\left|\frac{db}{d\Theta}\right|$

  where Θ = (plane) scattering angle

  b = distance of closest approach to nucleus → *effective* nuclear radius r

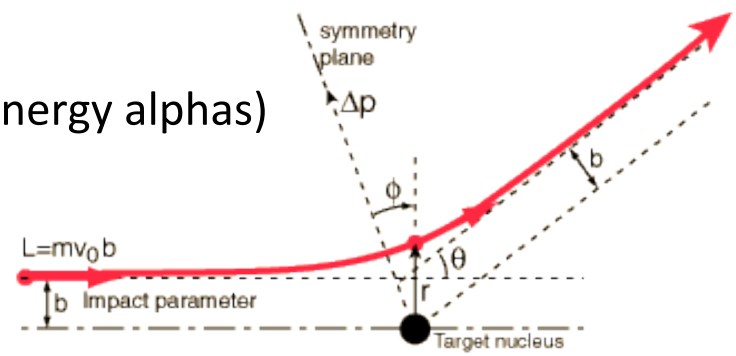
  Classical physics says: $b = \frac{1}{4\pi\epsilon_0} \cdot \frac{2q_1q_2}{mv^2}$ for electrostatic force,

  acting on projectile with charge q_1, mass m and speed v, near charge q_2.

  So $$\frac{d\sigma}{d\Omega} = \left(\frac{Z_1Z_2e^2}{8\pi\epsilon_0 m v_0^2}\right)^2 \csc^4\left(\frac{\Theta}{2}\right).$$

  Rutherford found b~ 3×10^{-14} m for gold

(actually: $r_{GOLD} \sim 7 \times 10^{-15}$ m -- he used low energy alphas)



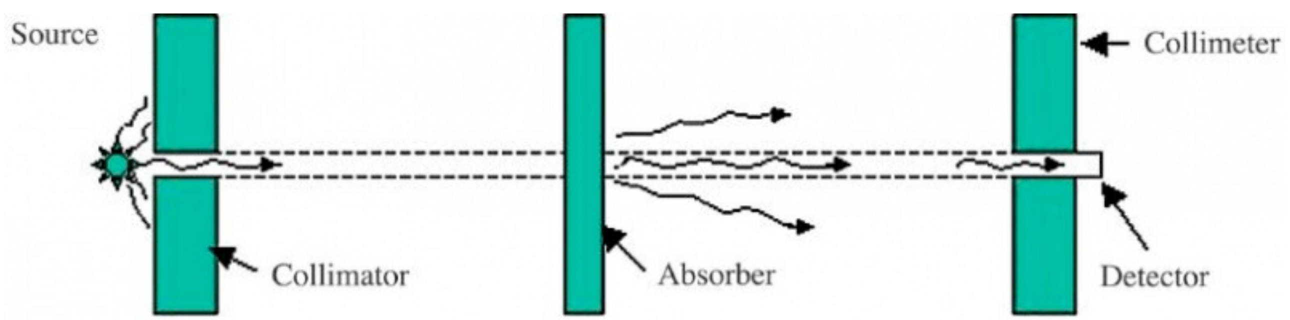
Attenuation and shielding for gammas

- "Good geometry" arrangement
  - Limited angular range for gammas from source
    - Scattered gammas are not counted: only survivors w/o interaction
    - Disappearance of gammas is simple sum of probabilities

      (photoelectric absorption) + (compton) + pair production
    - Can define linear attenuation coefficient μ: $I = I_0 \exp(- \mu t)$

I_0 = source intensity, t = absorber thickness, μ = attenuation coeff. (cm^{-1})

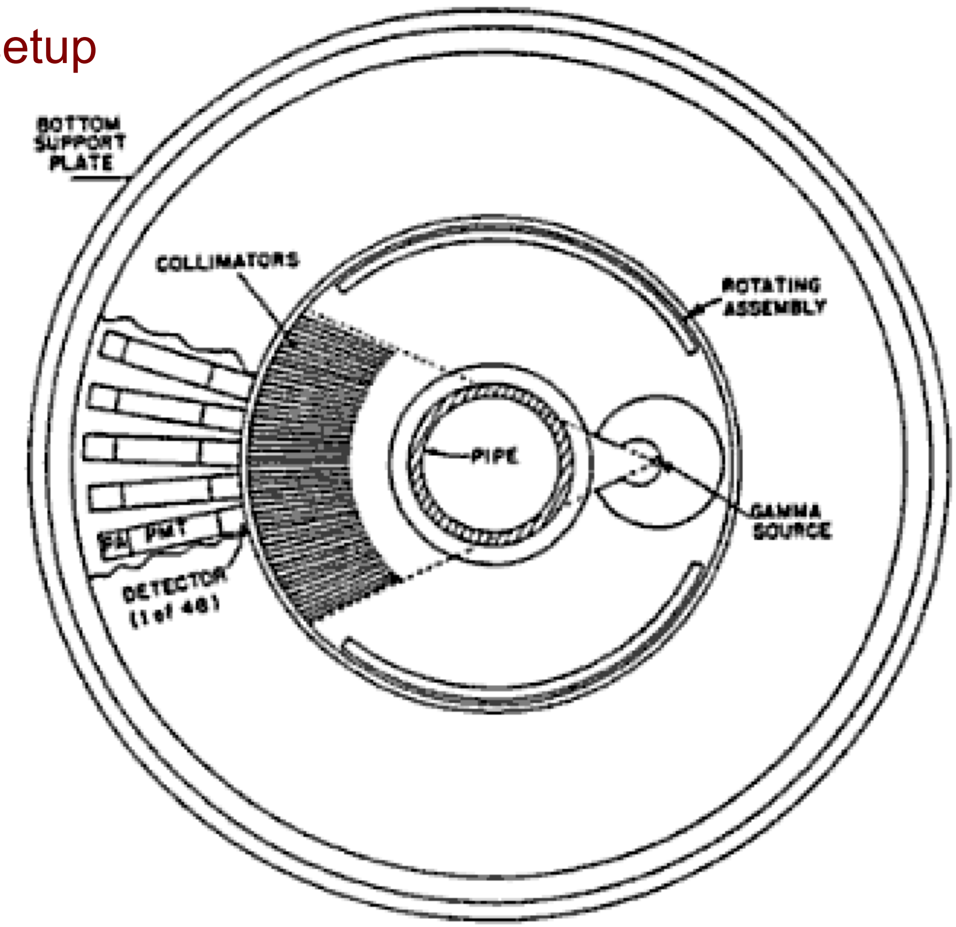
- Mass attenuation coeff = μ/(absorber density ρ) = cm/g^2
- For absorber = mixture of materials, use weighted sum of MAC's



Dwg from http://www.nucleonica.net/wiki/

Example: rotating scanner for inspecting pipe wall uniformity

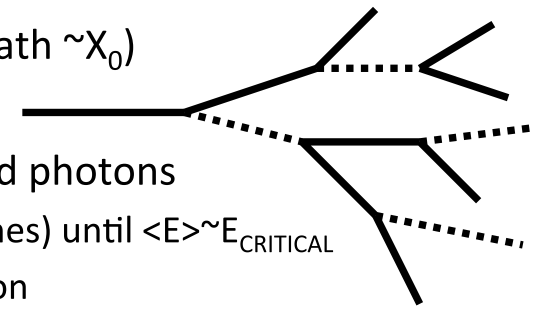
"good-geometry" setup

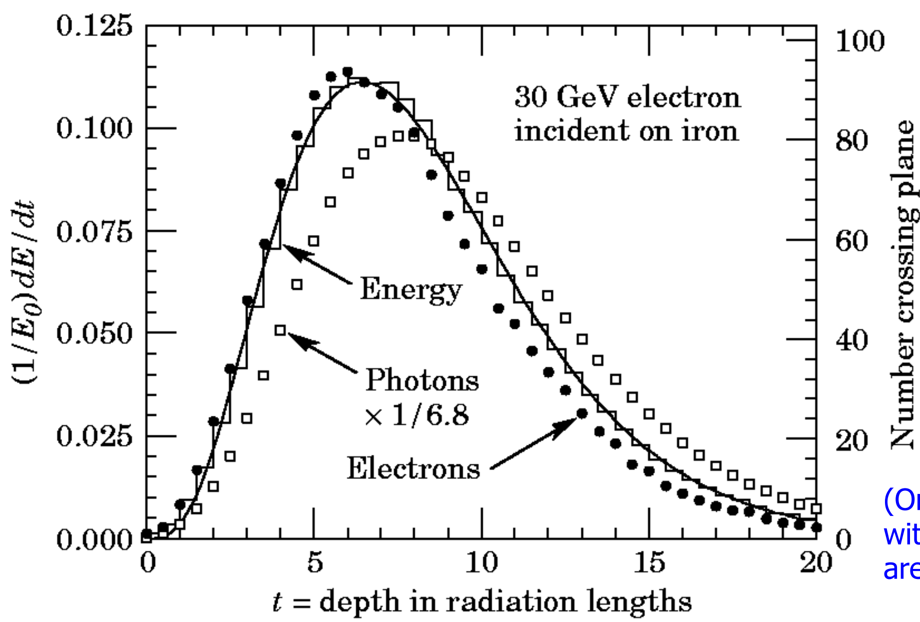


From: C. Melcher, *Industrial applications of scintillators*

Electromagnetic cascade development

- Electron entering dense matter soon brems (mean free path ~X_0)
- Brem photon soon *pair-produces* (mfp ~ (7/9)X_0)
  - etc, etc: result is a *cascade* or *shower* of electrons and photons
  - Number of particles builds up (and <E> per particle diminishes) until <E>~$E_{CRITICAL}$
    - Then brem losses become less important than ionization

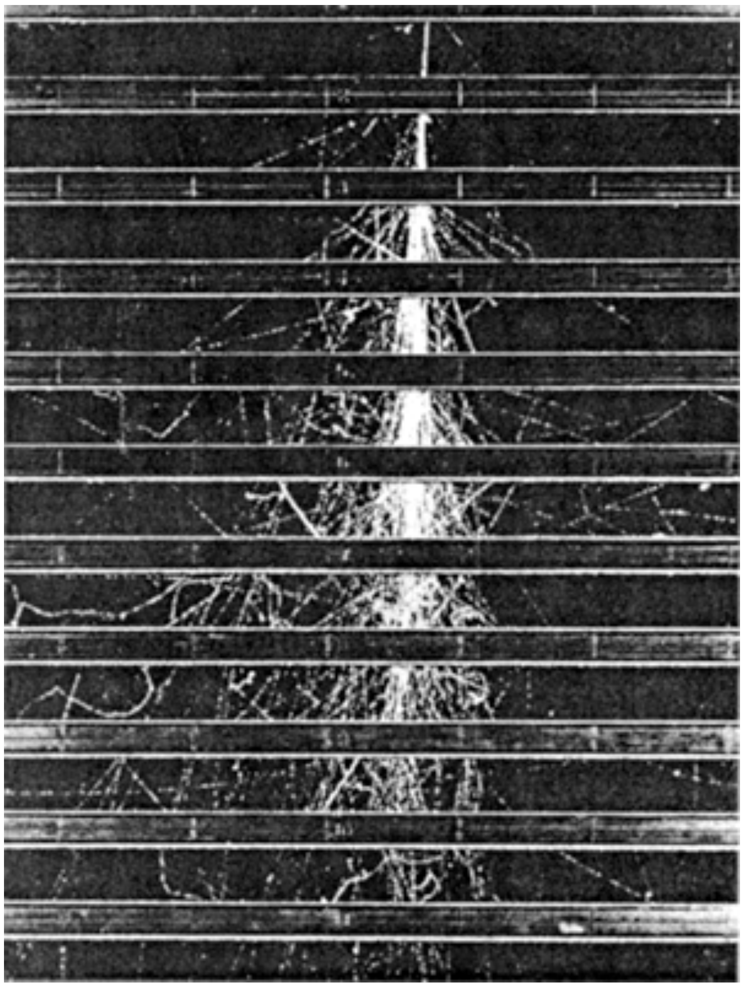




(Only particles with E>1.5 MeV are counted)

- Notice brem/PP cascade process does not *dissipate* energy, just swaps it from e's to photons and back again
- Main effect: divide energy up among more and more particles
- Cascade growth stops when average energy is too small to convert (less than E_C)
- Energy is lost to medium (heating) only via ionization, after e's drop below E_C

Electromagnetic cascade process



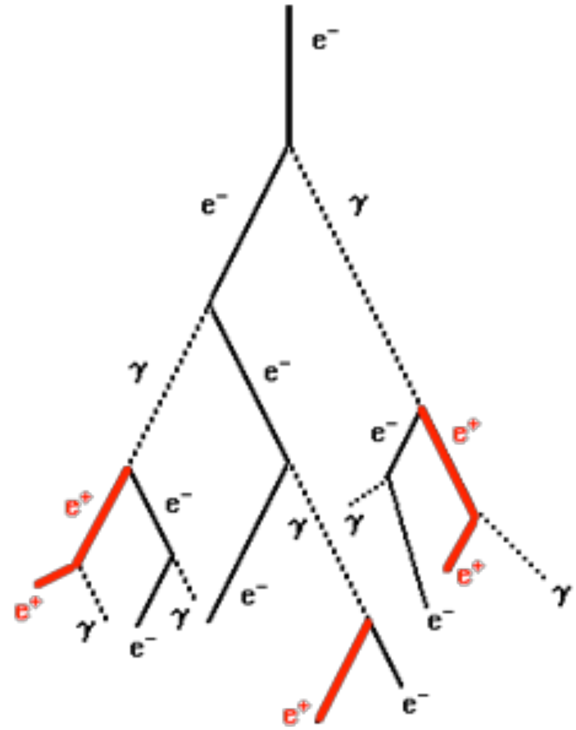


Photo from MIT cosmic ray group 1938

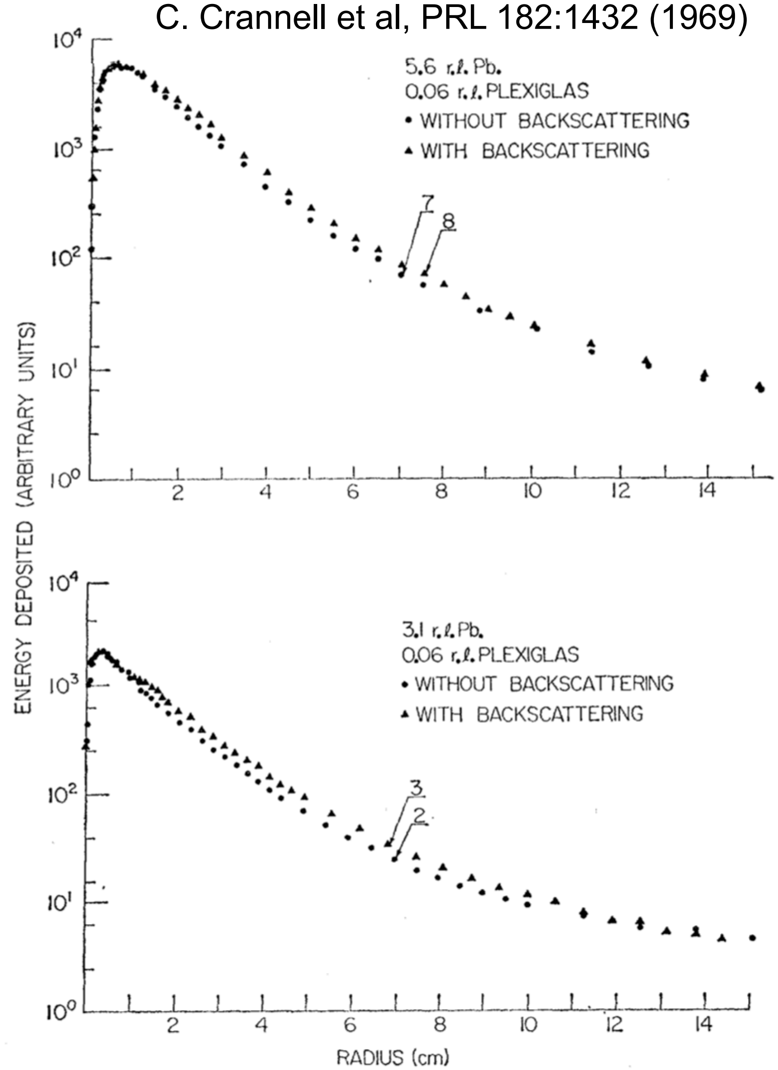
EM showers in detector design

- EM "calorimetry"
  - Basic way to measure energy of electrons or gammas
    - Thick absorber interleaved with particle counters
    - Number of particles vs depth and compare to calculations
      - can estimate total E of incoming e or gamma energy to O(10%)
  - Ideal = "Total absorption calorimeter" – nothing escapes
  - Real = truncated shower measurement with "punch-thru"
  - Tools
    - GEANT = particle physics industry standard for detector simulation
    - EGS = code developed at Stanford for cascade simulation
- Hadronic calorimeters
  - Same idea, but for protons/nuclei: much more complicated process! Much more depth needed.

Other cascade effects to consider

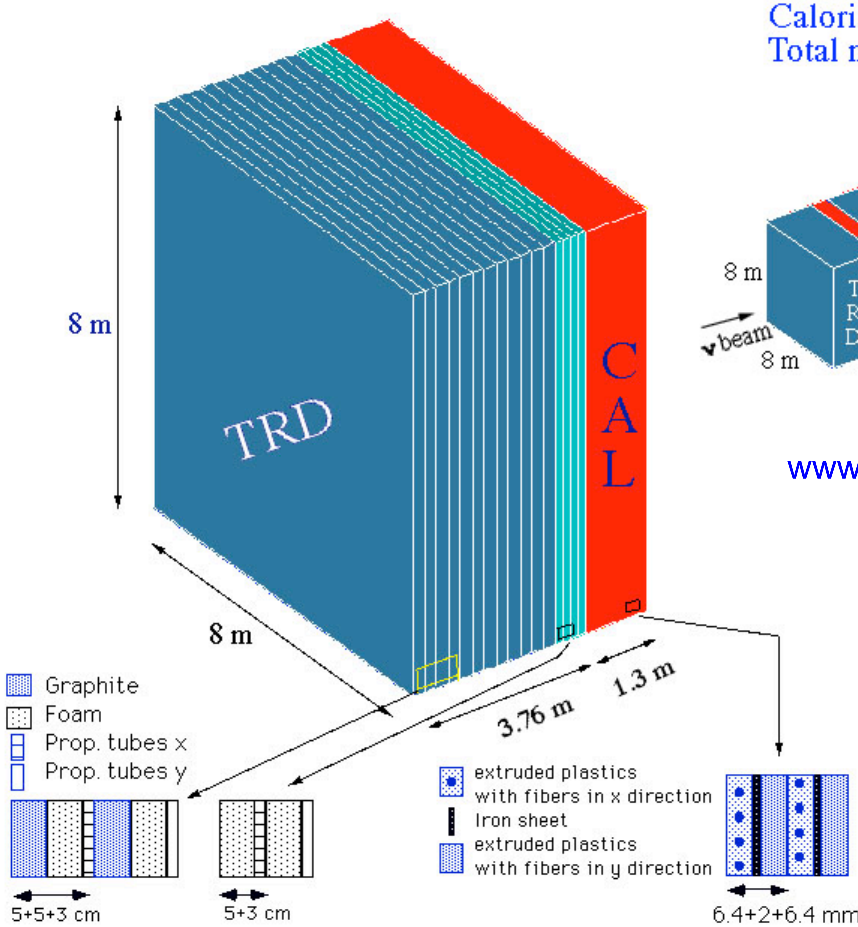
- Radial distribution of shower particles
  - Most particles in narrow core
- Transition effect:
  - Rapid change in relative populations of e's and gammas when Z changes
  - Transition radiation = x-ray photons produced at interface
  - Transition radiation detector (TRD) =
    - exploit TR to measure cascade
    - Very thin layers of Pb and plastic, use x-ray detector below many layers

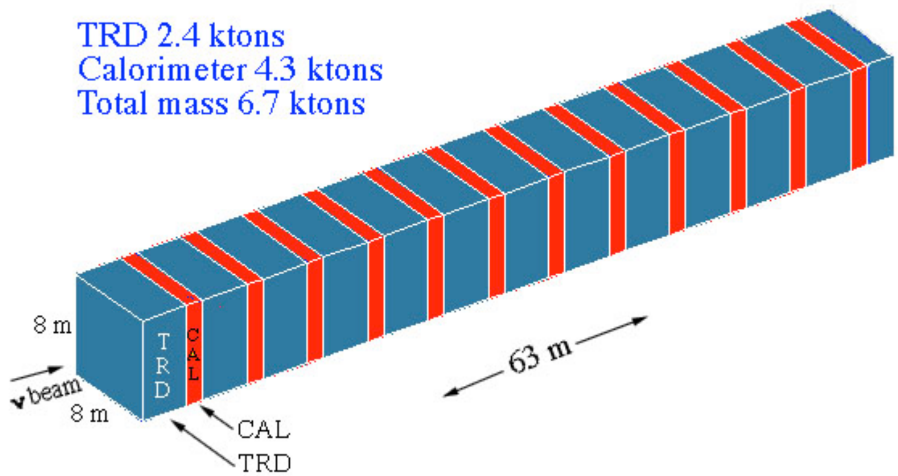
C. Crannell et al, PRL 182:1432 (1969)



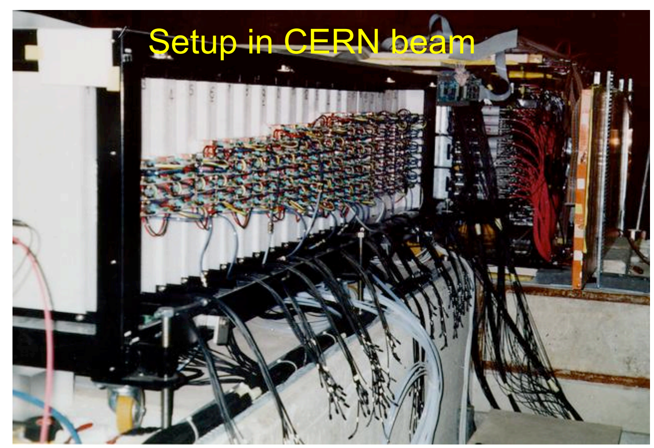
Transition radiation/calorimeter

- Example: NOE' detector (INFN/Naples, Italy)





www1.na.infn.it/wsubnucl/accel/noe/noe.html



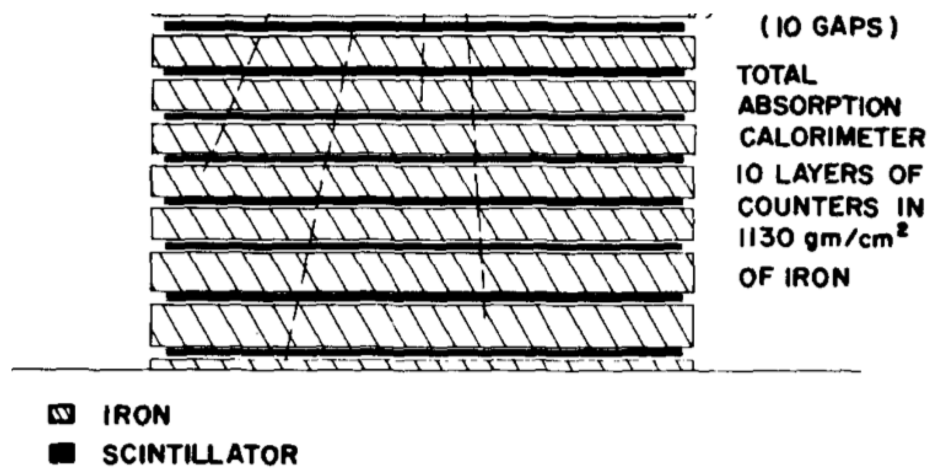
Example of calorimeter

- Total absorption calorimeter for protons
  - Proton interacts with nucleus producing mesons
    - Mesons interact again, or decay to electrons, photons, muons
  - Count charged particles present, at intervals in absorber
    - Area under plot of number vs depth = Estimate of total track length of shower particles in absorber
    - (Total track length)(energy loss per track per cm at critical energy) → estimate of total energy deposited by incoming proton

Underestimates E_0 due to:

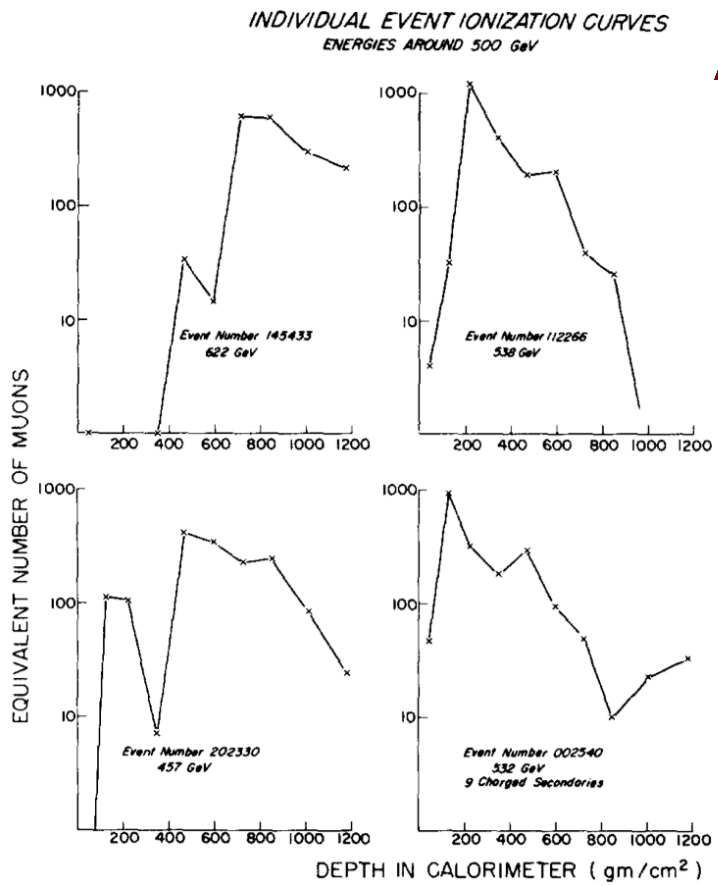
- "Punch-through" = remnant of shower that escapes through bottom
- Neutrons and neutrinos are not observed

Use simulation studies to estimate missed energy

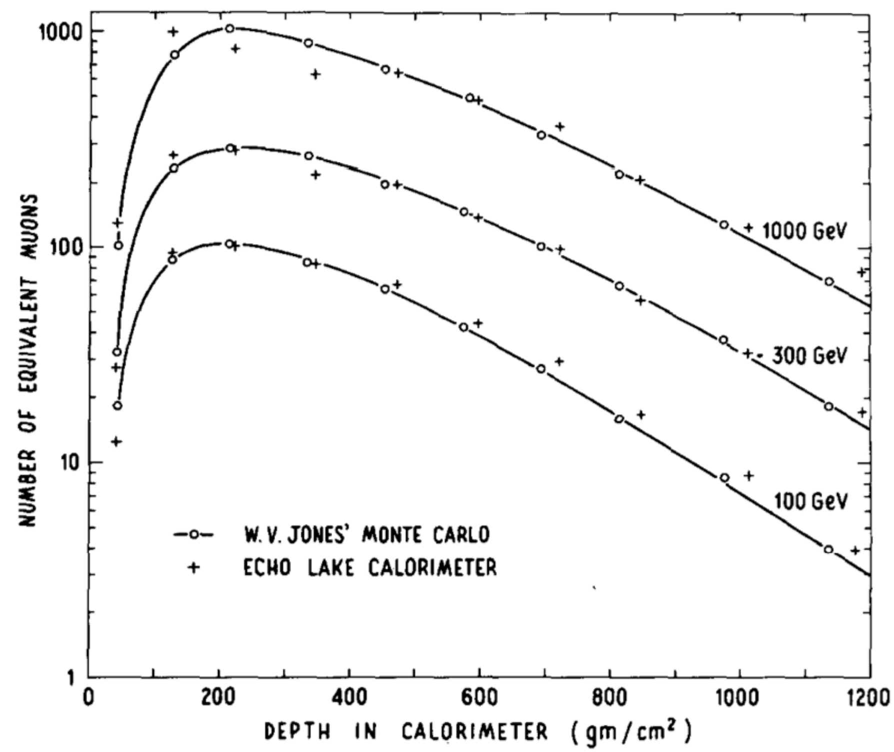


Individual showers and average showers

- Individual proton events show large fluctuations
  - Nuclear interactions inject many mesons, photon cascades have shorter length than overall hadron-electromagnetic cascade
- Average curves from many events + simulations can be used to calibrate



Area under curve → estimate of total track length T
Assume most tracks are around critical energy, then
$E_0 \sim T\, E_c$

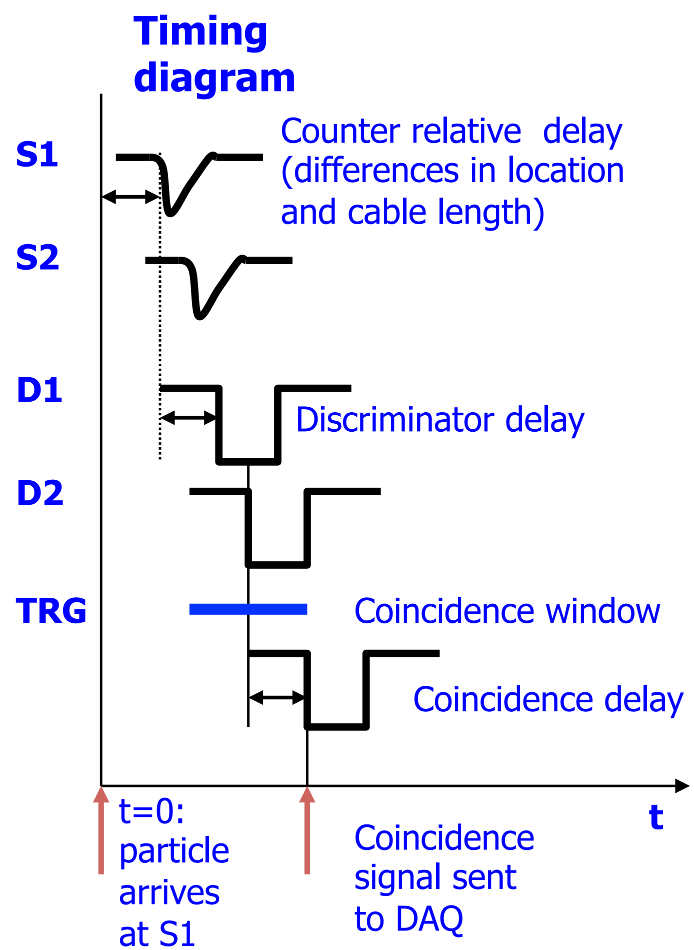


Counting experiments and statistics

- Coincidence measurements
  - Simultaneous signals from two or more detectors (eg scintillators) define events of interest
    - Must define what we mean by "simultaneous" (window width)
    - Must define "interesting" (coincidence pattern / logic)
  - Coincidence circuits (= logical .AND.)
    - Simple coincidence (channel 1 .AND. Ch 2 . AND. Ch 3…)
    - "Majority logic": any n-fold subset of all inputs
      - Eg, 2-fold for 3 inputs = (1 and 2) or (2 and 3) or (1 and 3)
    - Include possibility of veto (.NOT. or logical inversion) for inputs (Ch 1 . AND. Ch 2 . AND.(.NOT.Ch 3))
  - Must first convert raw PMT output to logic pulse: discriminator circtuit (one-shot)
    - Variety of analog pulses -> standardized logic signal
      - Specify voltage level (threshold) to trigger discriminator
    - Set output of the discriminator to a fixed height (choice of industry standard: TTL, NIM, etc), and duration (width)
- Q: how do we know counts are meaningful?

Delays and timing diagrams

- Discriminator output = coincidence circuit input
  - Must set threshold, width
- Must ensure signals from "simultaneous" detections arrive at the coincidence circuit at the same time
  - Time differences arise from
    - different transit times through the scintillator/light pipes/PMT, different cable lengths, etc
    - discriminator response times (leading edge of output vs input)
  - "Latest" signal defines the timing diagram
    - Add delay (cables or modules) to earlier signals as needed



Deadtime and Accidental coincidences

- Random signals from counters may fall within the coincidence time window and create an accidental count
  - Uncorrelated background particles happen to arrive simultaneously
  - Random noise from PMT, or ambient electronic noise
- To estimate the rate of accidental coincidences we need to know the resolving time of the system: *minimum time difference*
  - Depends on width of pulses input to the coincidence circuits, and the singles rate from each detector
  - Resolving time is measured by delaying one signal with respect to the other and plotting the coincidence counts per unit time (delay curve)
    - Set delay between counters to where delay curve is max
- If N_1 and N_2 are the singles rates and σ_t the resolving time, then the accidental rate will be approximately $N_A = \sigma_t N_1 N_2$
  - $\sigma_t N_1$ = time coincidence input 1 is on; N_2 = opportunities for an accidental
- Other factors affecting coincidence efficiency
  - Deadtime = time after detection when detector cannot generate a new signal
  - Jitter (random fluctuation) in timing of counter signal relative to particle arrival

Counting statistics

- Generally, we need to estimate probability of interesting events from a statistical sample of data:

Recall:

  - Statistic = single number, derived from data alone, describing some feature of a large data sample
  - Probability distribution = relative likelihood of different sample values

Some sample variables are integers (eg, counts); others are real-valued.

For $f(x)\,dx$ = Probability(x will be found in range $\{x \rightarrow x+dx\}$)

(with $f(x)$ properly *normalized* to give $\int_{\text{all } x} f(x)dx$ = probability of *any* x =1)

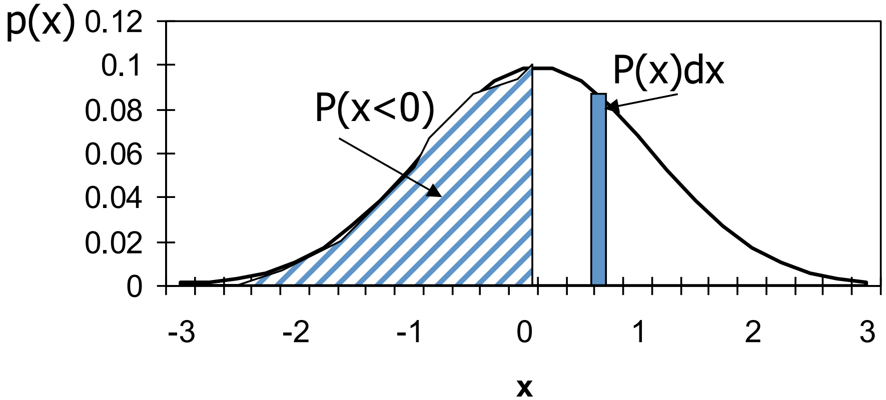
- $f(x)$ = *Probability Density Function* (PDF)
  - "differential probability distribution"
- $F(x) = \int_{-\infty}^{x} f(x)dx$ = *Cumulative* probability distribution
  - "integral distribution"

  $F(-\infty)=0$ and $F(+\infty)=1$; F is monotone increasing with x
- Estimate PDFs by making a *histogram* of experimental results (limited samples of x)

  histogram = bar graph of number of occurrences vs x

  $N(\Delta x) = N_{total}\, f(x)\, \Delta x \rightarrow f(x) \sim N(\Delta x)/(N_{total}\, \Delta x)$

- PDF p(x)=probability of x in range x' to x' +dx

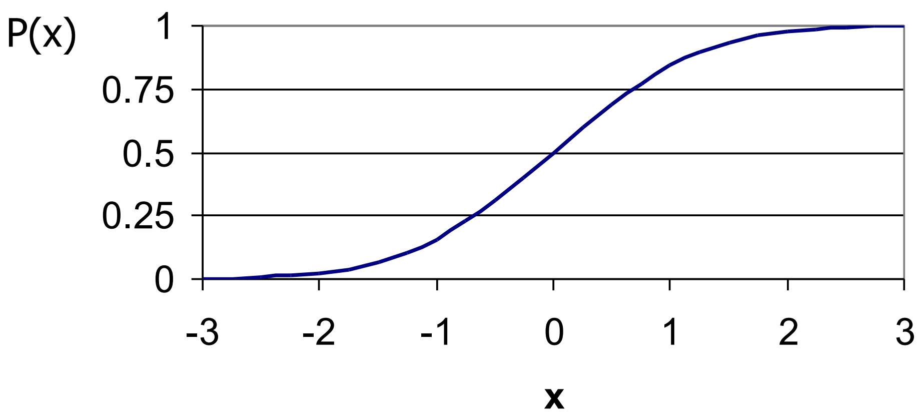


$$p(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

Normal (Gaussian) PDF

- "Probability distribution" P(x)=(*cumulative* or *integral* distribution) =probability of x<x'

$$P(x) = \int_{x_{MIN}}^{x} p(x)dx \quad \text{(where } x_{MIN} \text{ could be } -\infty\text{)}$$



Cumulative Standard Normal distribution

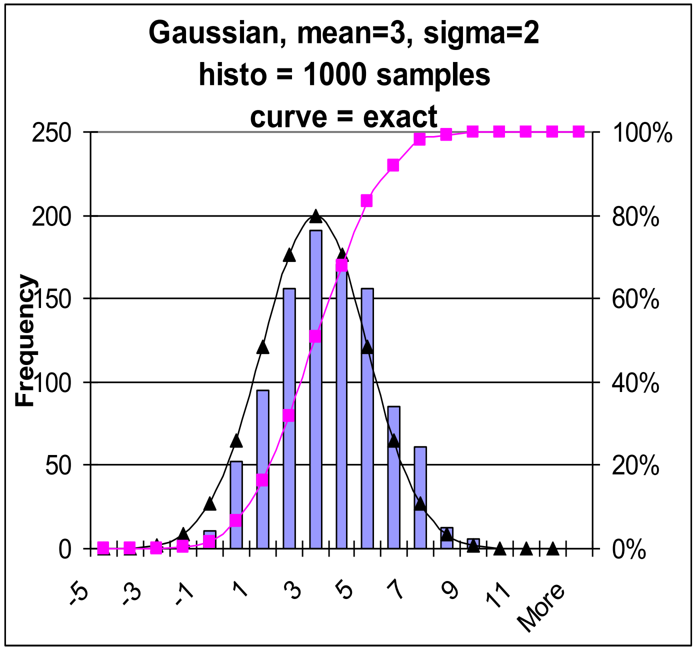
(erf(x)="error function")

$$P(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$$

Examples of distributions and histogram

- Table = list of data, binned in x
  - Simulated sample of 1000 events
  - Histogram = plot of table
- Curves = plots of underlying probability distributions
  - Black = PDF used to generate data sample: Gaussian centered on x=3
  - Pink = cumulative Gaussian probability (P(getting *any* value < x))

x	Frequency	Cumulative %
-5	0	0%
-4	0	0%
-3	1	0%
-2	2	0%
-1	11	1%
0	52	7%
1	95	16%
2	156	32%
3	191	51%
4	171	68%
5	156	84%
6	85	92%
7	61	98%
8	13	99%
9	6	100%
10	0	100%

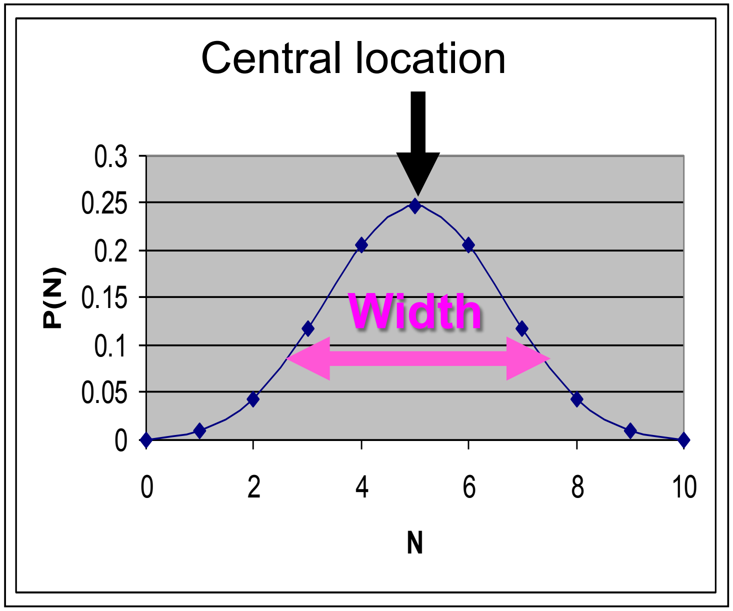


Descriptive parameters for PDFs

Commonly used *statistics*:

- Measures of central location:
  - mean $\langle x \rangle = \Sigma x_i / N$ (*sample* mean)
  - median = x at which F(x)=0.5
  - mode = x at which f(x)=maximum
  - for symmetrical distributions, mean=median
- Measures of width of distributions:
  - *variance* σ^2 (σ = *standard deviation*)
  - $\sigma^2 = \Sigma(x_i - \mu_1)^2 / N$
  - but μ_1 = mean of *true* PDF
  - we can only *estimate* μ_1 with $\langle x \rangle$
  - Best estimator for σ^2 is
  - $s^2 = \Sigma(x_i - \langle x \rangle)^2 / (N - 1)$ = *sample variance*
- Central moments:
  - *deviation* from mean $d_i = x_i - \langle x \rangle$
  - $\langle d^n \rangle = \mu_n$ = *nth central moment* (average of d_i^n)
  - $\mu_n = \int_{\text{all } x} (x - \mu_1)^n f(x)\, dx$

eg, n = 3 (skewness) gives measure of asymmetry, etc



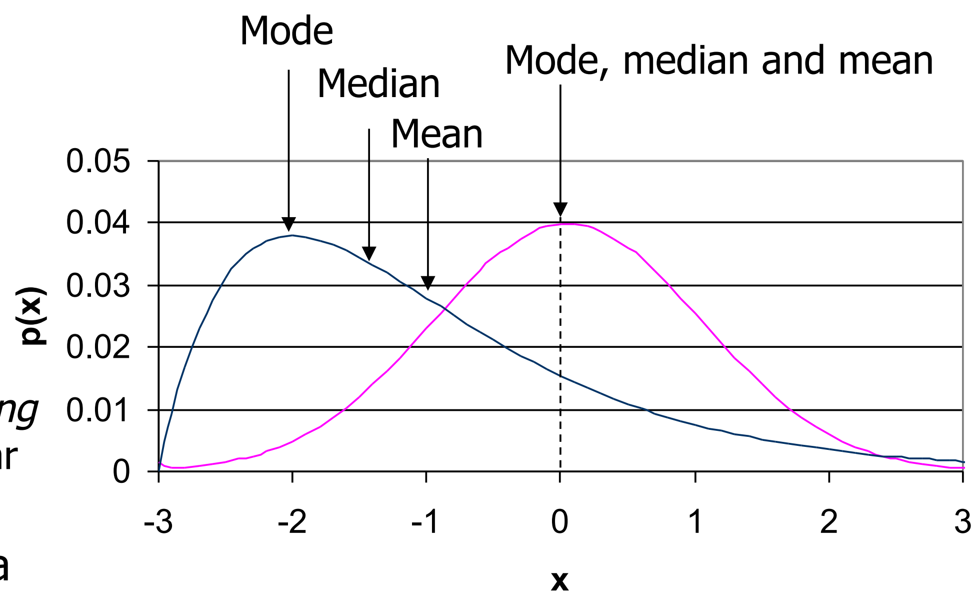
Examples of central location

Mean value (average): $\bar{x} = \frac{1}{N}\sum_{1}^{N} x_i$

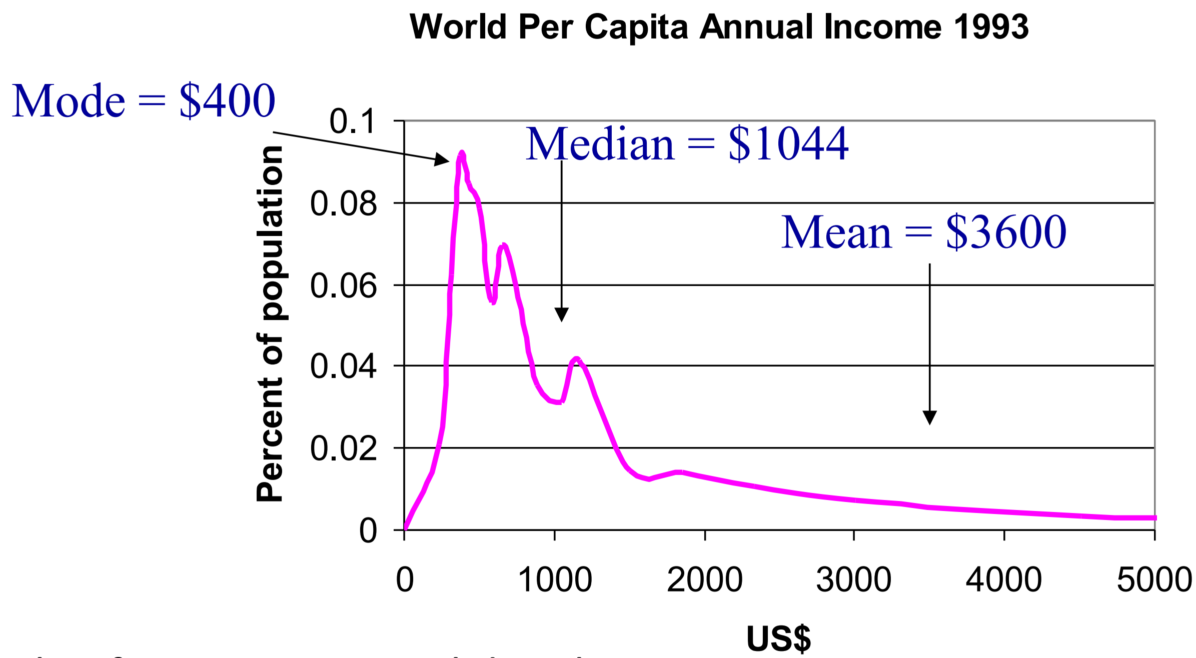
Median (center of sorted list): $x_{MED} = x_{\frac{N}{2}}$ in $sort_{\uparrow}(\{x_i\})$

Mode (peak value): x_i such that $P(x_i) = \max P(x)$

- Mean=median=mode for *symmetrical* distributions
- Median is more *robust* than mean for skewed PDFs
  - Mean is sensitive to *outliers* (few values far from central)
  - Median is sensitive only to *long tails* (significant population far from central value)
- Mode = *most likely* value for a *single* sample



Measures of central location



- Example of an asymmetrical distribution:
  World per capita income distribution (for 1993)
  - A small percentage of people have very large incomes, relatively
  - Mean = poor estimate of central value for highly skewed PDF
    - Long tail "pulls" the average value up

Measures of distribution width

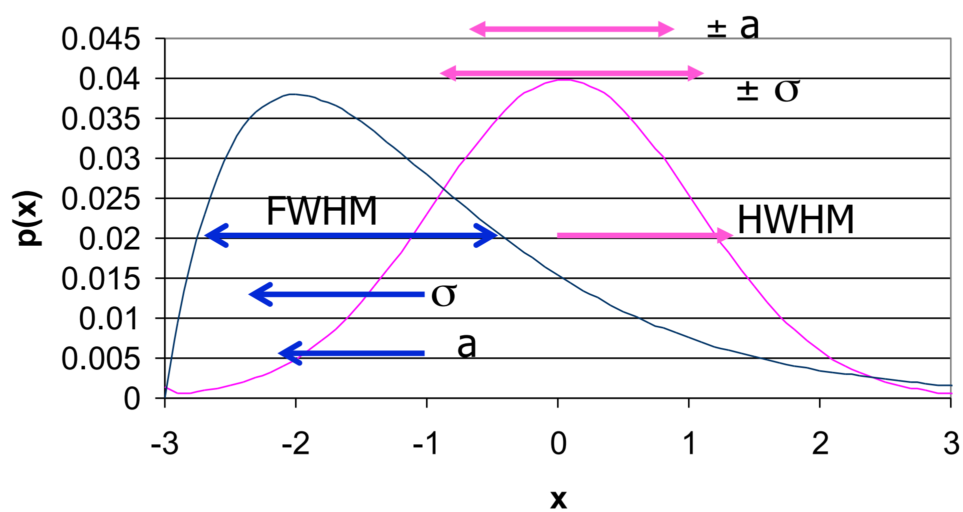
- Variance:

  $$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

  this is *sample variance* :

  Use N - 1 because we used

  the *same data set* to find $\bar{x}$



  - *Population variance*: Use N if we somehow know the mean value *a priori*
  - Note that mean and variance are 1st and 2nd *moments* of PDF
  - Variance has special significance in statistics (more later)
- Standard deviation: $\sigma=\sqrt{\sigma^2}$
  - Most commonly used measure of width
- Mean absolute deviation: $a = \frac{1}{N}\sum_{i=1}^{N}|x_i - \bar{x}|$
  - Not often useful
- Full- or Half-Width at half maximum (FWHM/HWHM)
  - Commonly used in engineering

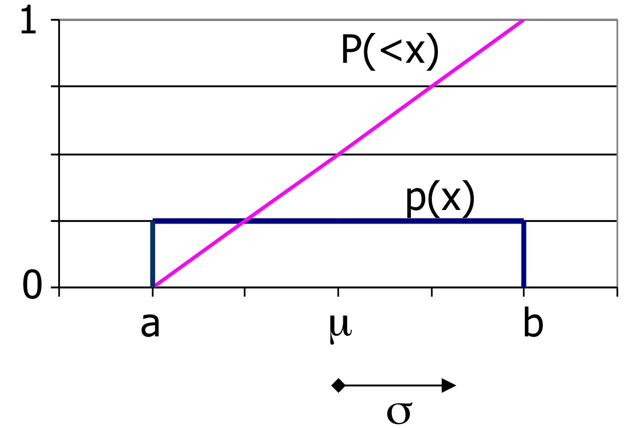
Famous probability distributions

- Uniform distribution
  - Only PDF generator available on all computer systems
    - Can construct all others from this (more later)

$$p(x) = \frac{1}{(b-a)} \qquad (a \le x \le b)$$

$$P(<x) = \int_a^x p(x)dx = \frac{(x-a)}{(b-a)}$$

$$\mu = \int_a^b x\,p(x)dx = \frac{1}{2}(a+b)$$



$$\sigma^2 = \int_a^b (x-\mu)^2\,p(x)dx = \frac{1}{12}(b-a)^2$$

e.g., for range 0 - 1: $\mu = 0.5 \qquad \sigma^2 = \frac{1}{12}$ (small!) skewness = 0, kurtosis = -1.2

- Binomial (Bernoulli) distribution
  - Describes experiments with *binary* outcome
    - Coin flips, win/lose lottery, detect/don't detect particle…

event x = success, $\bar{x}$ ("not x") = failure

are the only possible outcomes

n = number of trials

p = probability of success *per trial*

$$p_B(r; n, p) = \binom{n}{r} p^r (1-p)^{n-r}$$

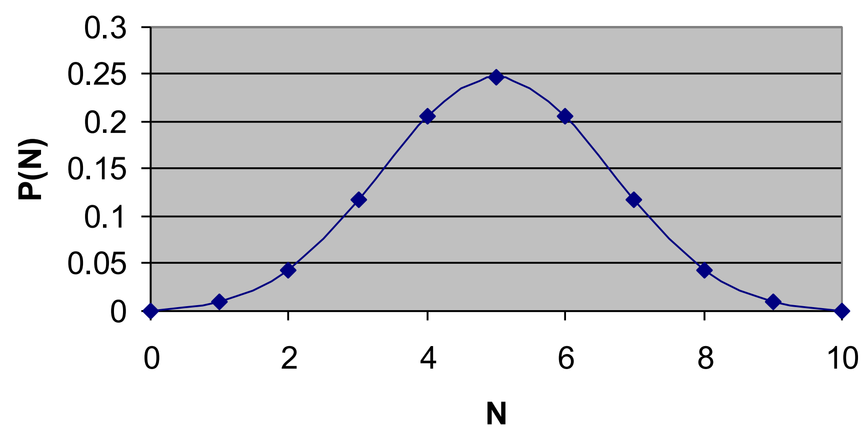
= probability of r successes in n trials

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

= # of combinations of n things taken r at a time (*order* of the r successes within the n trials is immaterial)

Binomial distribution for 10 flips of a coin, with probability 0.5 of getting heads on each flip.

N = number of heads observed in 10 flips
mean = 5
mode = 5
Width = $\pm$1.6



Applications of binomial distribution

Contents of a *single bin* in a histogram (n total entries)

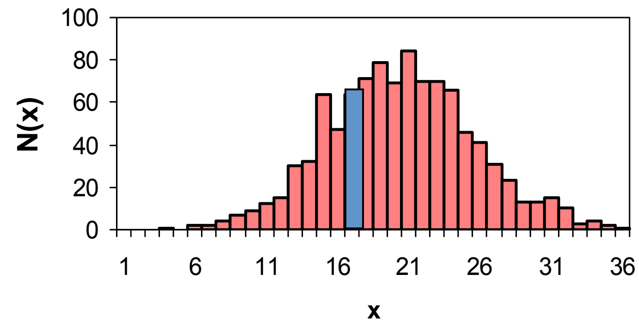
x = inside bin, $\bar{x}$ = outside $\qquad p$ = probability of x

From the data, we estimate $\hat{p} = \frac{r}{n}$, r = contents of bin

$$so\ q = (1 - \hat{p}) = 1 - \frac{r}{n}$$

$$\hat{\sigma}^2 = npq = n \cdot \hat{p} \cdot (1 - \hat{p}) = r\left(1 - \frac{r}{n}\right)$$

$$\mu \pm \sigma = r \pm \sqrt{r\left(1 - \frac{r}{n}\right)}$$



here: r=67, n=1000, so p=.066, σ=7.9

Note : Event counts are *integers,* but efficiencies are *real* numbers

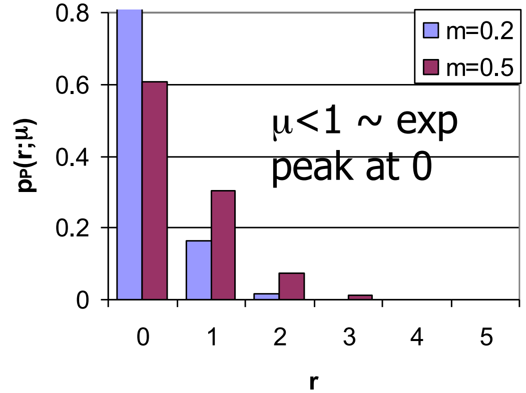
Change variable to *relative* frequency : $r \rightarrow \frac{r}{n}$

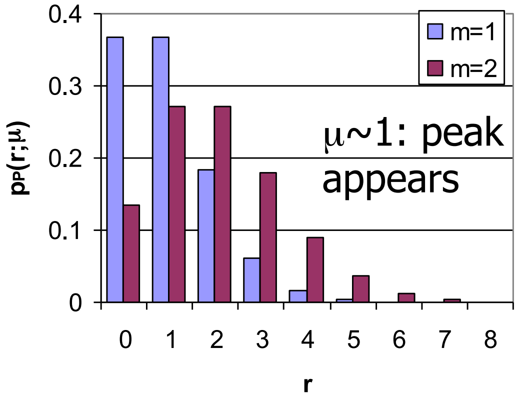
$$\mu \rightarrow \frac{\mu(p_B(r;n,p))}{n} = \frac{np}{n} = p, \quad \sigma^2 \rightarrow \frac{\sigma^2(p_B(r;n,p))}{n^2} = \frac{npq}{n^2} = \frac{p(1-p)}{n}$$

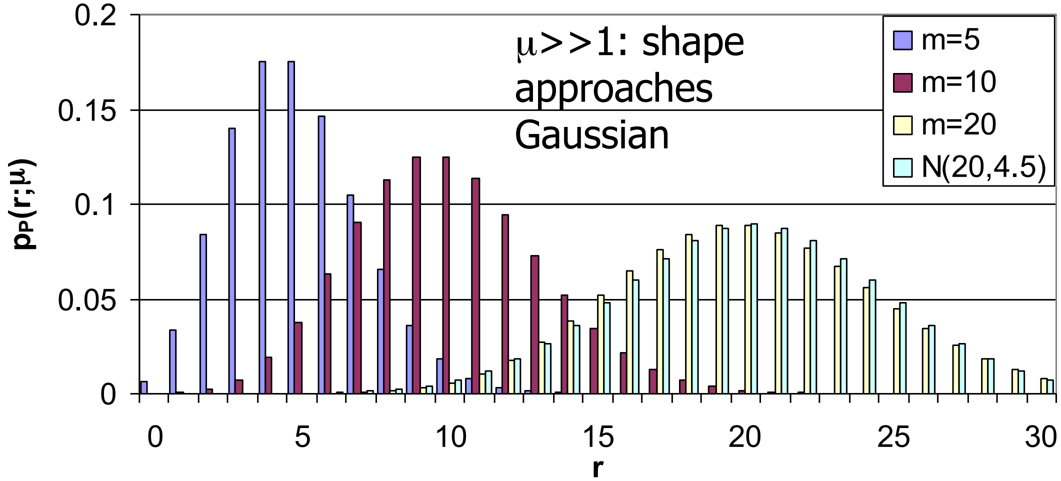
Common assumption: *Poisson limit* (for small p: more on this later) $\sigma = \sqrt{r}$

Poisson distribution

- Limiting case for binomial distribution with p small and n large
  - p → 0 and n → ∞ such that (np) ~ constant







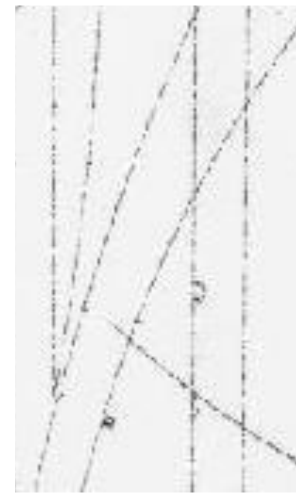
$$p_B(r;n,p) \Longrightarrow p_P(r;\mu) = \frac{1}{r!}\mu^r e^{-\mu}$$

$$\text{mean value} = \mu, \quad \sigma^2 = \mu \rightarrow \sigma = \sqrt{\mu} \quad \text{skewness} = \frac{1}{\sqrt{\mu}} \rightarrow \text{tail always to the right}$$

~ Symmetrical (→ normal distribution) for large μ;

Use p_P to approximate p_B for large n : $\mu = np$

Applications of Poisson distribution



- Notice $p_{POIS}(r)$ is limited to r = *integer* only
  - Value of μ (not necessarily integer) should be "small"
  - Range of r is bounded on the left (by r=0)
  - Approaches normal dist. for μ "large" (far from 0)
  - Has only one parameter: μ
- Applications
  - Radioactive decay counting data: μ = mean counts/sec

  Then prob. of r counts/sec is $p_P(r;\mu) = \frac{1}{r!}\mu^r e^{-\mu}$

  Example: suppose μ = 20 counts/sec

  Then prob of 30 counts in any one sec period is about 1% (~ same for 10 counts)

  - Bubbles in a bubble chamber track: μ = mean bubbles/cm: • • • • • • • •

  (or hits in any ionization tracking detector: ionization is proportional to Q^2)

  Then prob. of getting r bubbles/cm is given by Poisson distribution

  Example: suppose proton (Q=1) tracks have 9 bubbles/cm on average

  Searching 1000 pictures, a track is found with 1 bubble/cm -- a quark (Q=1/3)?

  Prob that this is just a random fluctuation of a proton track? p_p (20;9)=0.0011
  so expect to see this with odds ~1:1000…

  Not unlikely enough! ("Discovery Threshold" is currently around $P\sim10^{-4}$)

Poisson assumptions

- Physical situations where *Poisson Assumptions* are valid lead to behavior reflecting the exponential and Poisson distributions:
  1. p(1 event) in interval δx is *proportional to* δx: p=g δx
  2. Occurrence of an event in some interval δx_j is *independent* of events (or absence of events) in any other non-overlapping interval δx_k
  3. For sufficiently small δx, there can be *at most 1 event* in δx
  – Examples: ionization in a gas, goals scored in a soccer match, requests for documents on a web server, radioactive decays
- From these we can derive the exponential and Poisson distributions:

Prob of 1 bubble in δx : $p_1(\delta x) = g\delta x \quad (\mathit{from}\ \#1)$

Prob of 0 bubbles in δx : $p_0(\delta x) = 1 - p_1 = 1 - g\delta x \quad (\mathit{from}\ \#3)$

$$p_0(x+\delta x) = p_0(x) \bullet p_0(\delta x) = p_0(x)(1 - g\delta x) \quad (\mathit{from}\ \#2)$$

$$\therefore \frac{p_0(x+\delta x) - p_0(x)}{\delta x} = -g\,p_0(x) \rightarrow \frac{dp_0}{dx} = -g\,p_0$$

$$\mathit{Solution:} \quad p_0(x) = e^{-gx}$$

Prob of exactly r bubbles in $x + \delta x$:

$$p_r(x+\delta x) = p_r(x) \bullet p_0(\delta x) + p_{r-1}(x) \bullet p_1(\delta x) \quad (\mathit{from}\ \#3)$$

$$\therefore \frac{p_r(x+\delta x) - p_r(x)}{\delta x} \rightarrow \frac{dp_r}{dx} = -g\,p_r(x) + g\,p_{r-1}(x)$$

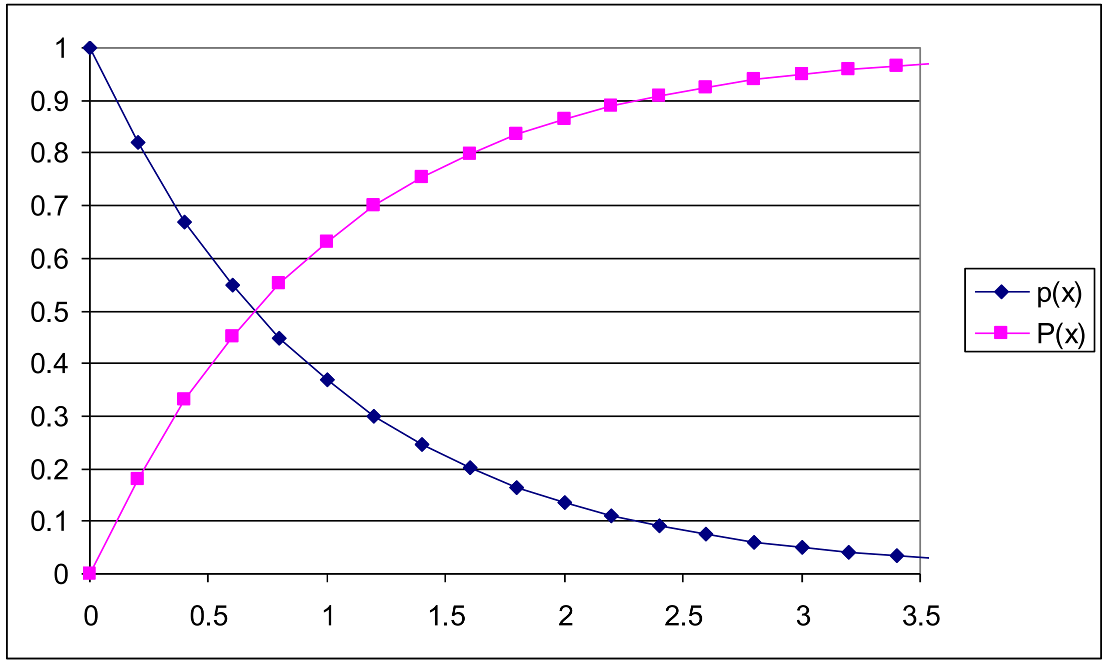
$$\mathit{Solution:} \quad p_r(x) = \frac{1}{r!}(gx)^r e^{-gx} = \text{Poisson} \ (\mu = gx)$$

So *exponential* distribution = *gap length* distribution $p_0(x)$ between events in a Poisson process (gaps in a bubble chamber track or ionization trail)

Exponential distribution

- Special case of frequent interest: probability of getting *exactly one* event under Poisson case:

$$p_P(1;\mu) = \mu^1 e^{-\mu}$$



Example: radioactive particle has mean life λ

Probability that particle decays within any time window of length t is given by cumulative exponential distribution (integral distribution)

$$PDF: \quad p(t;\lambda) = \frac{1}{\lambda} e^{-t/\lambda}$$

$$Cumulative: \quad P(<t;\lambda) = 1 - \frac{1}{\lambda} e^{-t/\lambda}$$

Gaussian ("Normal") probability density fn (PDF)

- Gaussian = famous "bell-shaped curve"
  - Describes IQ scores, number of ants in a colony of a given species, wear profile on old stone stairs...

  All these are cases where:
  - deviation from norm is equally probable in either direction
  - Variable is continuous (or large enough integer to look continuous - far from the "wall" at n = zero)
- *Real-valued* PDF: $f(x) \rightarrow \; -\infty < x < +\infty$

  $n(x;\mu,\sigma) = (1/\text{sqrt}[2\pi\sigma^2]) \exp[-(x-\mu)^2/2\sigma^2]$
- 2 independent parameters: μ, σ (central location and width)
- Properties:

  Symmetrical, mode at μ, median=mean=mode

  Inflection points at $\pm\sigma$

  Standard normalized form:

  scale x by σ, shift origin to μ

  $n(x;0,1) = (1/\text{sqrt}[2\pi]) \exp[-x^2]$

  Cumulative distribution:

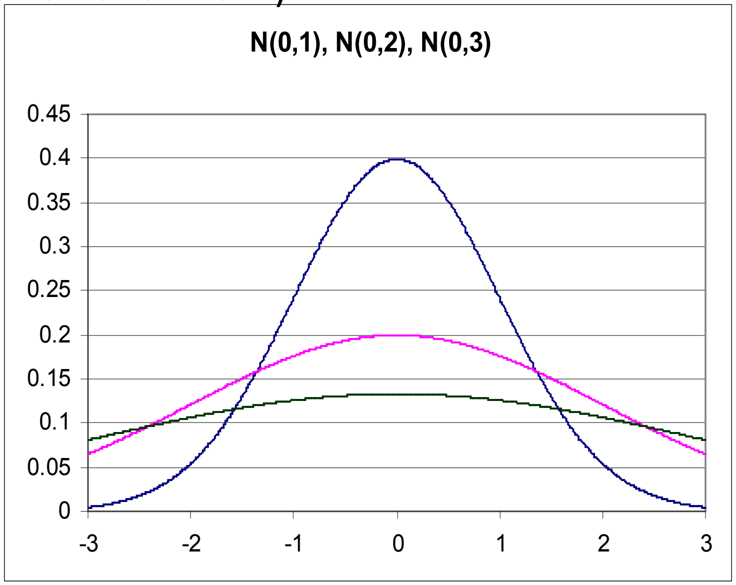
  $N(x) = \int_{-\infty}^{x} n(x;0,1)dx = \text{erf}(x)$

  Area within (prob. of observing event within)

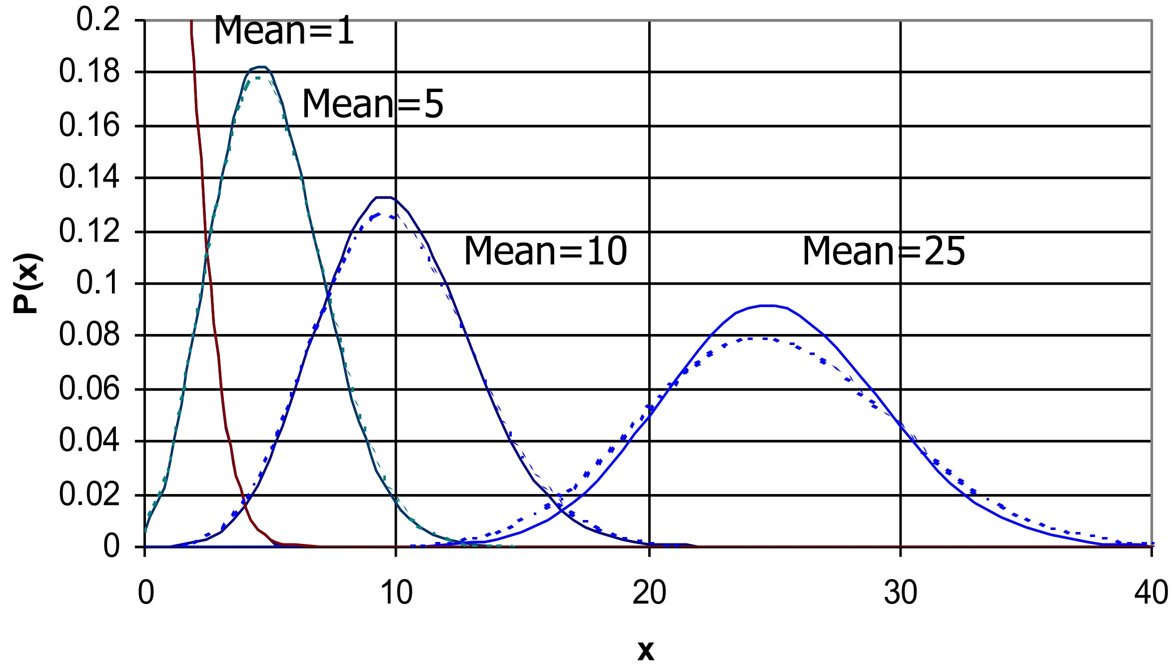
  $\pm 1\sigma = 0.683$ = erf(1)-erf(-1)

  $\pm 2\sigma = 0.955$ = erf(2)-erf(-2)

  $\pm 3\sigma = 0.997$ = erf(3)-erf(-3)



Binomial, Poisson, Normal



Shown above:

- Binomial for 100 trials, p=0.01, 0.05, 0.10, 0.25 (solid)
- Poisson for μ = 1, 5, 10, 25 (dashed line)

Poisson is broader and has peak slightly below μ
Both become similar to Gaussian N(μ, $\sigma=\sqrt{\mu}$) as mean value gets larger
(Gaussian for μ=25, σ=5 is indistinguishable from Poisson on this scale)

Significance of normal distribution

- *Central Limit Theorem:*

  "Given N independent random variables x_k, each with μ_k and σ_k specified (but *not* details of individual PDF's), the random variable $z = \Sigma x_k$ has mean value $\mu = \Sigma \mu_k$ and variance $\sigma^2 = \Sigma \sigma_k^2$ and for large statistics, its PDF will be Gaussian, ie

  $(\Sigma x_k - \Sigma \mu_k) / \text{sqrt}[\Sigma \sigma_k^2] = n(x;0,1)$"

- Applies to: any situation with real-valued result where several *independent* processes *add:* additive errors.
- Parameters μ,σ are independent (and converse: if a random variable has μ,σ independent, it is normal).

  Given N random numbers x_k from a normal distribution,

  the sample mean $\mu = (1/N)\Sigma x_k$

  and sample variance $s^2 = \Sigma \sigma_k^2 / (N-1)$

  are *independent statistics*

Application examples:

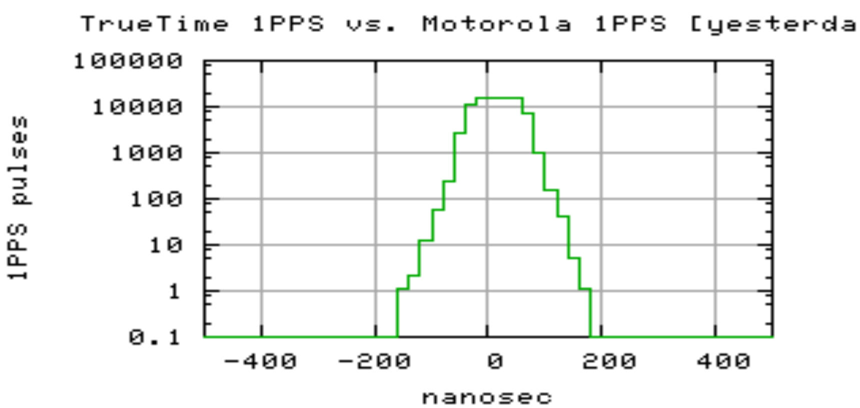
- Random walk of 100 steps. Each step is independent of others, any probability distribution for direction and length of each step (but μ, σ^2 known).
- To make a simple Gaussian random number generator, just take sum of 12 uniformly distributed numbers on [0,1):

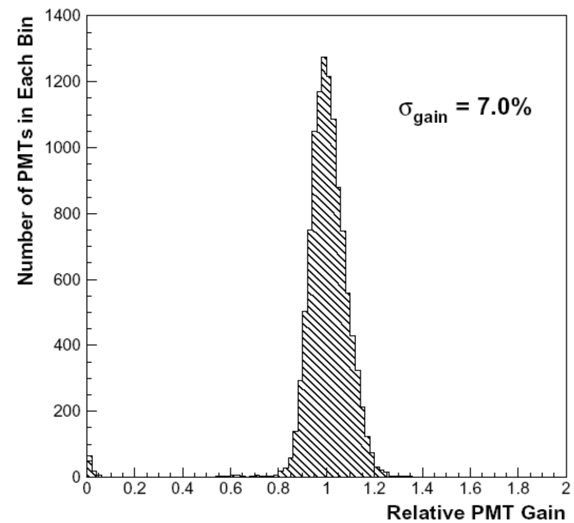
  $x=\Sigma (u_k - 6)$; x will be distributed ~ $n(x;0,1)$

  (recall: *uniform*(0;1) has $\mu= 0.5$, $\sigma^2= 1/12$)

Examples of normal distributions

- Gain distribution for 11,000 photomultiplier tubes which are supposed to be ~identical (from Super-K neutrino experiment)
  - Gain depends on many independent factors (tube manufacture batch, geometry, power supply stability, cable characteristics...)





- GPS clock times (again from Super-K, yesterday's data)
  - Time fit depends on many indep. factors (exact antenna location, software design, receiver characteristics...)

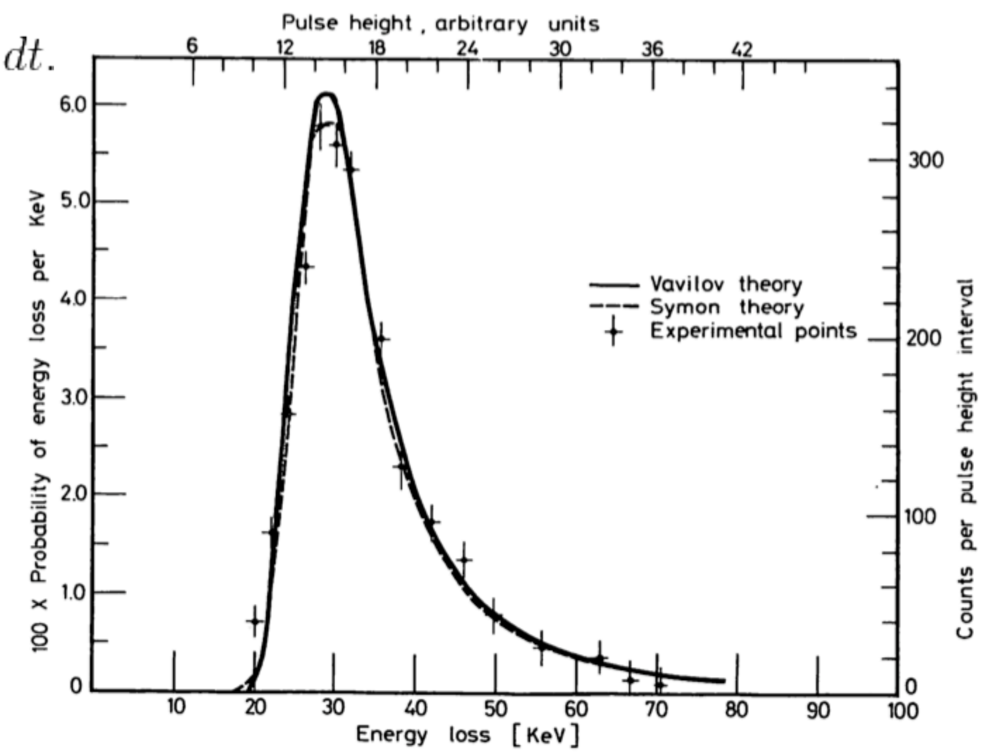
Special physics PDFs: Landau distribution

- Landau described the distribution of energy losses for particles passing through a thin layer of absorber
  - Long tail on the right, cut off sharply on the left -> most losses near average, but big losses possible, and less infrequent than for Gaussian
  - Related to the Cauchy (Breit-Wigner) and Gaussian distributions

$$p(x) = \frac{1}{\pi} \int_0^\infty e^{-t \log t - xt} \sin(\pi t)\, dt.$$

- Landau is a member of the Stable Distribution family, which includes Gaussian, Cauchy and Delta functions

See math.uah.edu/stat/special/Stable.html



Breit-Wigner distribution

- B-W distribution describes distribution of energy seen in decays of very short-lived states (resonances)

In particle physics units, where ħ = c = 1:

E = center-of-mass energy of the decay products,

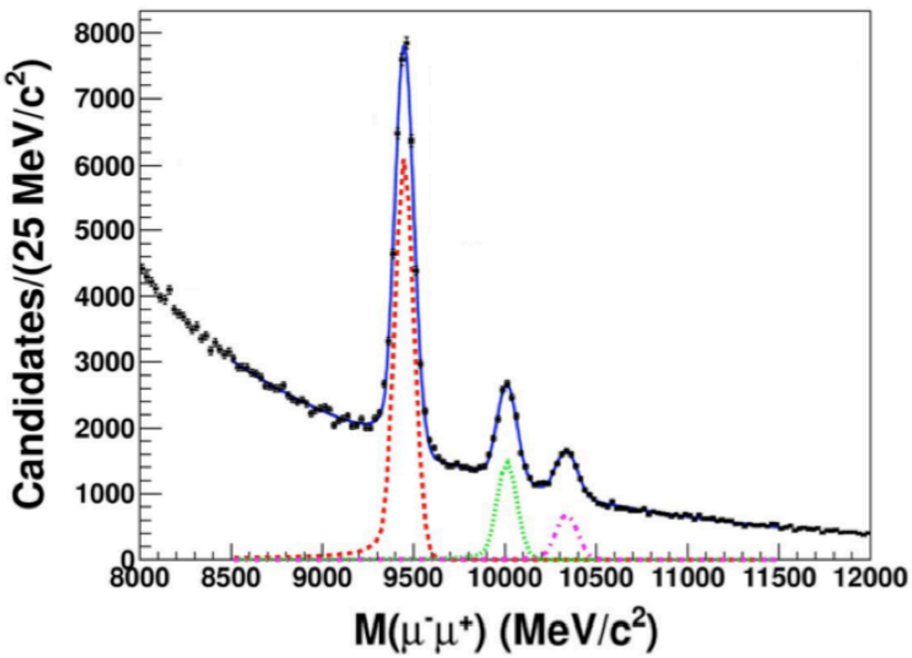
M is the mass of the resonance (peak location),

Γ is the resonance width, and mean lifetime τ = 1/Γ.

$$f(E) = \frac{k}{\left(E^2 - M^2\right)^2 + M^2\Gamma^2} \ ,$$

$$k = \frac{2\sqrt{2}M\Gamma\gamma}{\pi\sqrt{M^2+\gamma}}$$

$$\gamma = \sqrt{M^2\left(M^2+\Gamma^2\right)} \ .$$



Example from CERN LHCb experiment, showing Upsilon meson (= b-anti-b quark pair) and 2 excited states, decaying promptly to muons.
B-W shapes are fitted to excess over smooth random background counts in each bin
lhcb-public.web.cern.ch

Application of statistics: hypothesis testing

- Given
  - 2 data samples: $\{x_{i=1...N}\}$, $\{x_{j=1...M}\}$ OR
  - data sample $\{x_{i=1...N}\}$ and model $f(x;\underline{\theta})$
- We want to test the "Null Hypothesis":

H_0 : $\{x_{i=1...N}\}$ and $\{x_{j=1...M}\}$ are drawn from the *same* population distribution

[alternatively: $\{x_{i=1...N}\}$ is drawn from $f(x;\theta)$]

  - "non-parametric tests" = no assumptions made about the underlying population distribution

Chi-squared test (Pearson's test)

  - Histogram the $\{x_i\}$ to estimate differential dist.
  - n_i = number of entries in bin i ($x_i \leq x < x_i+\delta x$), δx=histogram bin width
  - To test f(x):

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{n_i - m_i}{\sqrt{m_i}} \right)^2 \quad where \quad m_i = \int_{x_i}^{x_i+\delta x} f(x)\,dx$$

(m_i is not necessarily integer)

Note: must exclude empty bins ($\sqrt{0}$ in denominator)

Denominator is Poisson estimate of σ_f^2

Central limit theorem $\rightarrow \chi^2$ is correct for large n_i or large N_{BINS}

Chi-squared

Sampling distributions describe statistics of data samples as a whole

- Chi-squared distribution $\chi^2 \equiv \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{\sigma^2}$ summed deviations squared, in units of σ^2

$$p(\chi^2;\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}(\chi^2)^{\nu/2-1}\exp(-\chi^2/2)$$

$$= \chi^2 \text{ PDF for } \nu \text{ degrees of freedom}$$

ν = number of *independent* variables in sum

  – For example: if μ is unknown a priori, we must use *average* x as *estimator* for μ:

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{\sigma^2} \rightarrow p(\chi^2;\nu = N-1)$$

Also, note that $\sum_{i=1}^{N}(x_i - \bar{x})^2 = (n-1)s^2$

so $z = \frac{(n-1)s^2}{\sigma^2}$ is also a χ^2 variable,

with PDF given by $p(\chi^2;\nu = N-1)$

If $u_1, u_2 \ldots u_M$ are χ^2 variables with different ν_i, then

$w = \sum_{i=1}^{M} u_i$ is a χ^2 variable following $p(\chi^2;\nu = \sum_{1}^{M}\nu_i)$

Chi-squared distribution

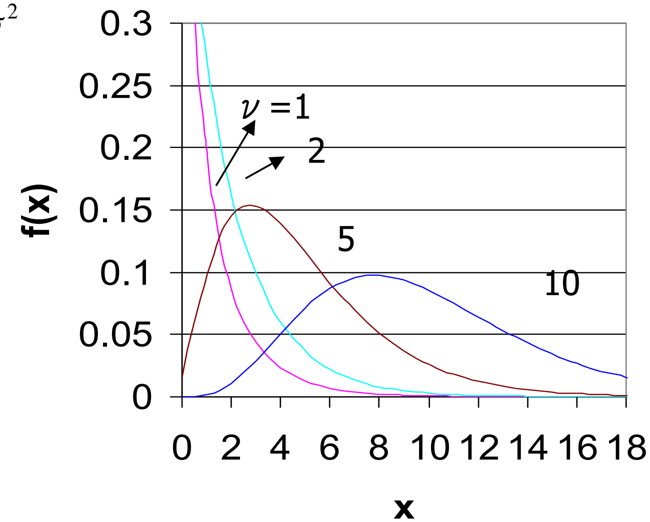
$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(x_i-\mu)^2}{\sigma^2},\quad \text{sum of deviations squared, in units of } \sigma^2$$

$$p(\chi^2;\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}(\chi^2)^{\nu/2-1}\exp(-\chi^2/2)$$

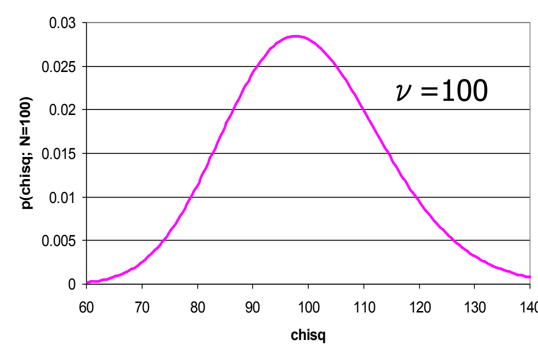
$$= \chi^2 \text{ PDF for } \nu \text{ degrees of freedom}$$

ν = number of *independent* variables in sum

Example: if we average N data points to estimate μ, ν=N-1



- Chisq distribution is
  - Monotone decreasing for $\nu \le 2$
  - Peaks at ν-2 for ν> 2
  - Has mean=ν, σ^2=2ν and → N(ν,2ν) for ν -> ∞



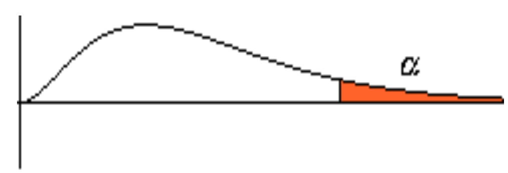
Integral distribution: $P(\chi^2_\alpha;\nu) = \int_0^{\chi^2_\alpha} p(\chi^2;\nu)\,d\chi^2 = 1-\alpha$

So $\chi^2 > \chi^2_\alpha$ occurs with probability $= \alpha$

Use to test for N(μ,σ^2) behavior:

Example: test hypothesis that $\{x_i\}$ come from N(μ,σ^2)

Then we should have $\chi^2 \equiv \sum_{i=1}^{N} \frac{(x_i-\bar{x})^2}{\sigma^2} \le \chi^2_\alpha(\nu = N-1)$



to have *confidence level* α in our hypothesis

Rule of thumb: for $\nu \ge 10$, $\chi^2_{0.5} \cong \nu \rightarrow \frac{\chi^2}{\nu} \approx 1$ is 50% probable

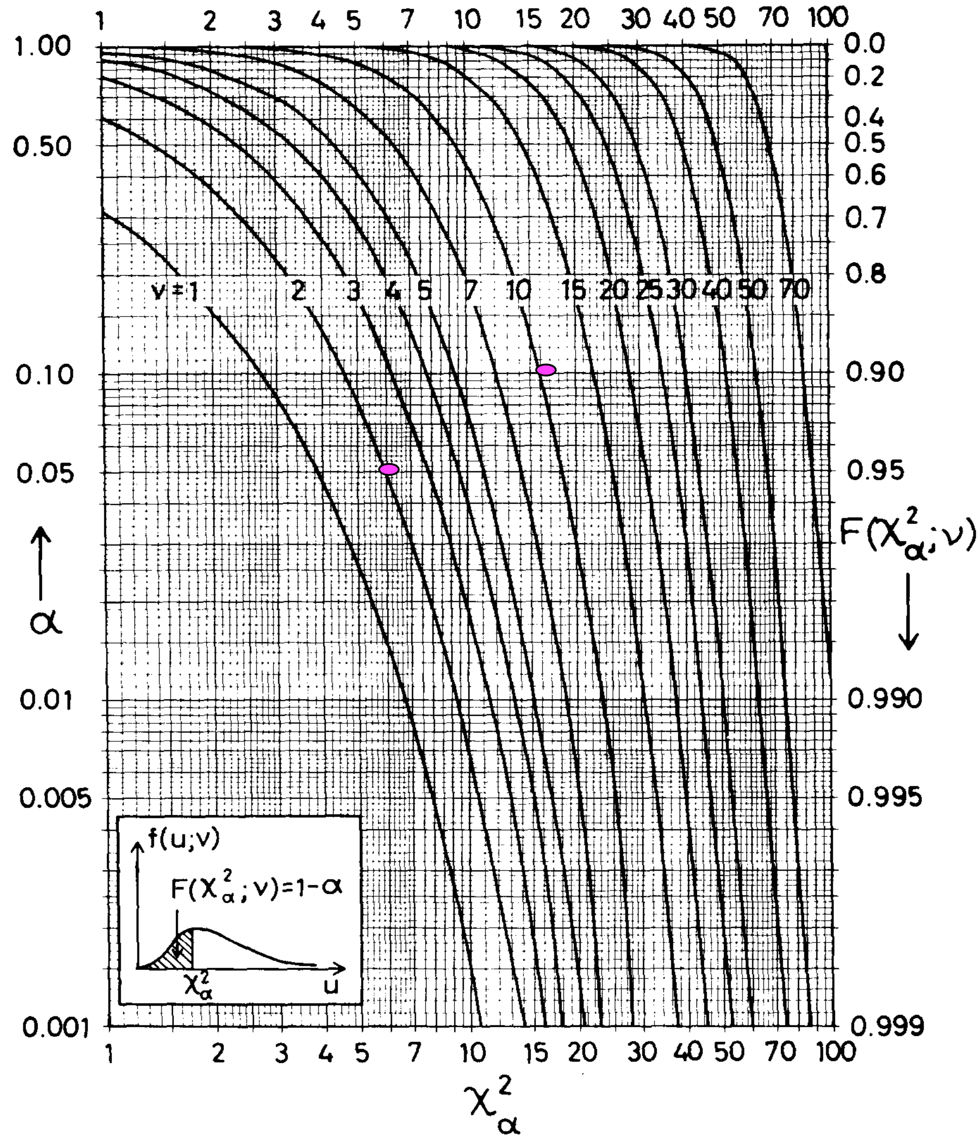
Nomogram for percentage points of χ^2

Use to find α given χ^2 and DOF, or χ^2 given α and DOF.

Examples ( ):

for 10 DOF, α = 10%→ χ^2=16;

For 2 DOF and χ^2 = 6 → α= 5%



(From Frodesen et al)