X-Ray Notes, Part I

X-ray Imaging

Images are characterized by the interaction of x-ray photons and tissue.

Physics

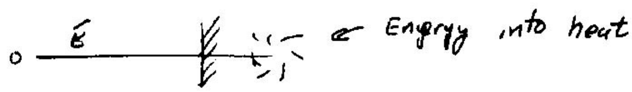
Definition: Radiation – a stream of particles or photons.

Particles: α ($^{2+}$He), e^- (electrons), β (electrons emitted from nuclei), β^+ (positrons), p^+ (proton), n^0 (neutrons)

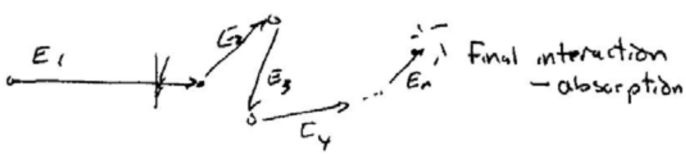
Photons: x-ray, γ, annihilation photons, etc.

Models for interaction of radiation and matter:

1. Absorption (generally low kinetic energy (KE))



2. Scattering



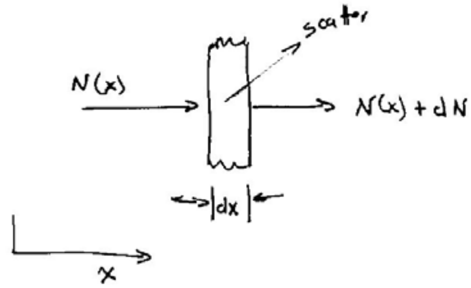
3. Not a typical interaction – a gradual loss of energy

The charged particles above (α, e^-, β, β^+, p^+) interact very strongly with tissue and typically do not pass completely through the human body and thus cannot be used for imaging. Of the above particles photons and neutrons(n^0) pass through the body with an appropriate amount of interaction for imaging (too little is also bad).

Behavior of Radiation Along a Line

Assumptions:

1. Matter consists of discrete particles separated by distances that are large compared to the size of the particles.
2. For a given path length along a line, an x-ray photon either interacts (with prob. *p*) or it doesn't and all interactions are independent.
3. Scattered photons scatter at a different angle and don't contribute to the continuing flux of photons along the line.



The change in the number of photons is:

$$dN \propto -N(x)dx$$

$$dN = -\mu N(x)dx$$

$$\frac{dN}{dx} = -\mu N(x)$$

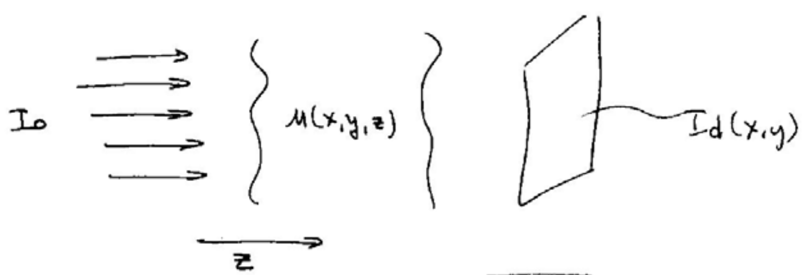
$$N(x) = N(0)\exp\left(-\int_0^x \mu(x')dx'\right)$$

were μ is the "linear attenuation coefficient" and has units (distance)$^{-1}$. For a constant μ:

$$N(x) = N(0)\exp(-\mu x)$$

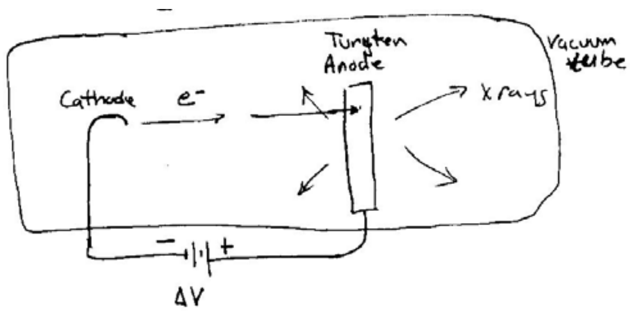
The Basic X-ray Imaging System

Now consider a parallel ray x-ray flux that has intensity I_0 (intensity is photons/unit area/unit time) the passes through a 3D object having a distribution of attenuation coefficients $\mu(x,y,z)$ and projects to an image $I_d(x,y)$:



$$I_d(x,y) = I_0 \exp\left(-\int \mu(x,y,z)dz\right)$$

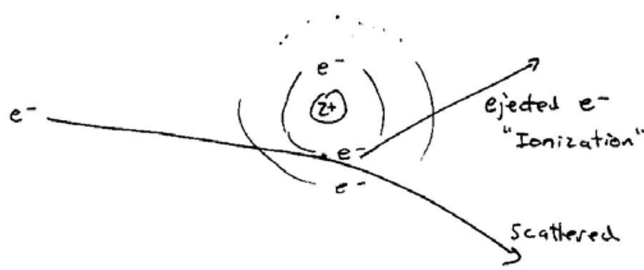
Generation of x-rays



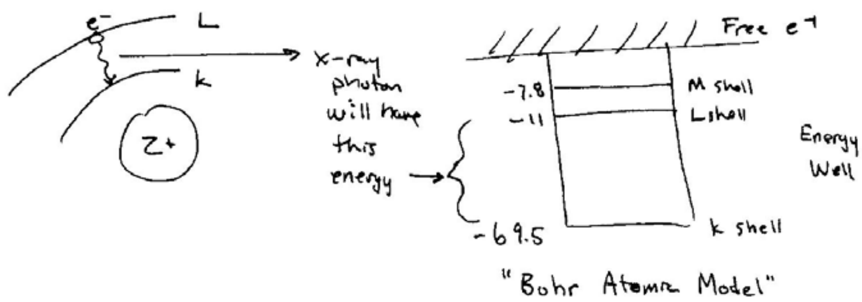
- Target is usually a high-Z, heavy element – typically W, tungsten.
- Electrons are accelerated by the voltage between the cathode and the anode.
- A potential energy of E=qΔv (e.g. e * 150 kV = 150 keV) all gets converted to kinitic energy E = ½ $m_e v^2$ (e.g. also 150 keV).

Kinds of electron interactions:

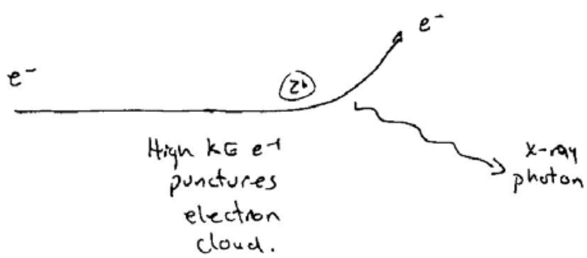
a. Inelastic (energy absorbing) scattering with atomic electrons – the ejection of a bound electron followed by emission of a photons from spontaneous energy state transitions.



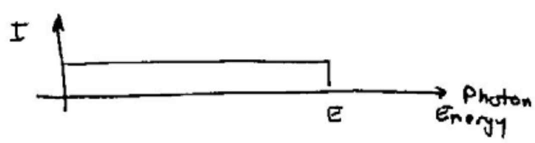
The Bohr model accounts for absorption/generation of discrete valued energies.
58.5 keV is one "characteristic" x-ray for W. Any combination of shell transition energies will also be characteristic energies (e.g. 3.2 and 61.7 keV). Very low energies are hard to observe due to other absorption processes.



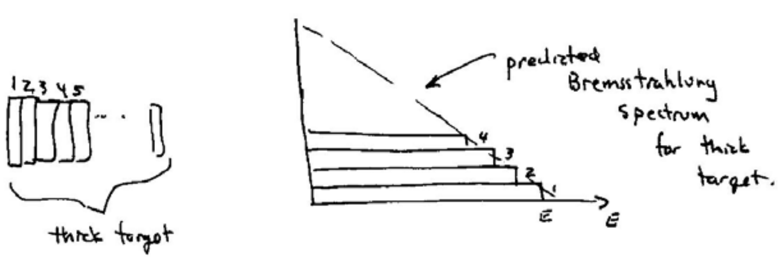
b. Bremsstrahlung "Braking" Radiation – Acceleration (change in direction) of electron by Coulomb attraction to the large, positively charged nucleus leads to the generation of photons (acceleration of any charged particle will do this).



For electrons of a particular energy, E, striking an infinitely thin target, Bremsstrahlung radiation will have a uniform distribution of energy between 0 and E.

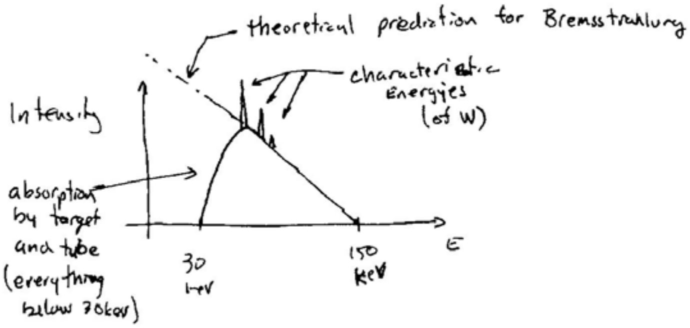


We assume that all electrons interact. For a thick target, it is often modeled as a series of thin targets where the highest energy impinging upon subsequent stages is reduced by the interactions. Each thin target produces a new uniform spectrum, but with a lower peak energy. The resultant spectrum is approximately linear from a peak at 0 keV to 0 at E.



The x-ray Spectrum

- For electrons with energy E, the maximum x-ray photon energy is E.
- $E = h\upsilon = \frac{hc}{\lambda}$
- Very low energy photons are absorbed by the target and by the glass in the x-ray tube.
- Spectrum will have a combination of Bohr (discrete) energies and Bremsstrahlung radiation:

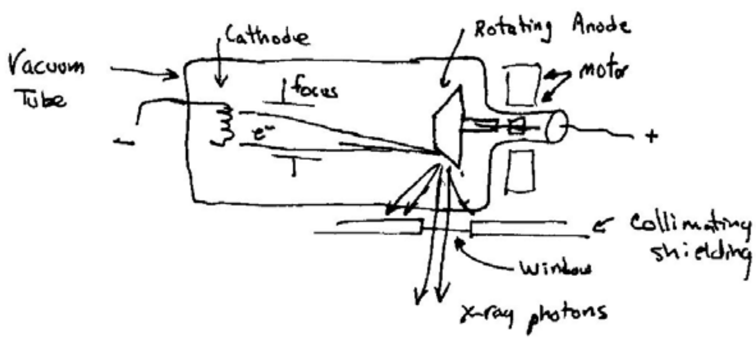


- The x-ray spectrum is function of photon energy: $I_0 = I_0(E)$
- I now represents energy/unit time/unit area or power/unit area.

Practical x-ray tube

Why Tungsten?

- x-ray spectrum in desired range
- High Z (high efficiency in stopping electrons)
- High melting point (3300 deg. C) – typical operation temp is ~2500 deg. C – this is due to the low efficiency of the electron to x-ray conversion (~0.8%). The rest goes into heat.
- Example:



- Rotation of target to reduce peak temp

- Shielding to collimate beam
- Window further filters x-ray spectrum ("hardens beam" – makes it have a higher average E)

The Attenuation Coefficient

We say above that the x-ray spectrum is a function of photon energy E: $I_0 = I_0$(E). The attenuation function is also a function of E: $\mu = \mu(x,y,z,\text{E})$. The new expression for the intensity at the output will not be:

$$I_d(x,y) = \int_E I_0(E)\exp\left(-\int \mu(x,y,z,E)dz\right)dE$$

Note: I_d tells us nothing about z or E – it only gives us x,y information.

The x-ray attenuation coefficient μ is, of course, also a function of material properties. Two of the most important properties that affect the attenuation coefficient are tissue density, ρ, and the atomic number Z. As most x-ray photon/tissue interactions are photon/electron interactions both ρ and Z will influence μ.

For x-ray photons, there are 4 main types of interactions (listed in order of increasing likelihood with increasing photon energy, E):

1. Rayleigh-Thompson Scattering
2. Photoelectric Absorption
3. Compton Scattering
4. Pair Production

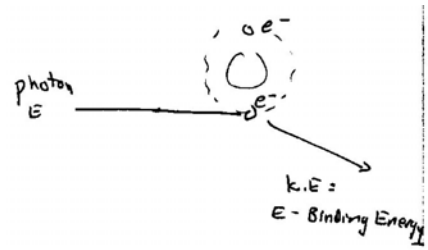
In general, we can write an expression for the attenuation coefficient as the some of these constiuent parts:

$$\mu(E) = \mu_{rt}(E) + \mu_{pe}(E) + \mu_{cs}(E) + \mu_{pp}(E) + \ldots$$

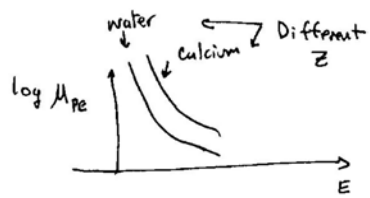
1. Rayleigh-Thompson Scattering or "coherent" scattering – atomic absorption with spontaneous emission at the same energy E. This is the same effect as is seen in x-ray

diffraction in crystals. This term is rarely important in the diagnostic energy range (50-200 keV).

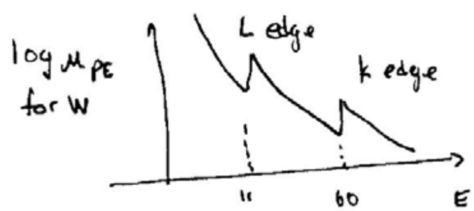
2. Photoelectric Absorption – Absorption of photon to ionize and eject an atomic electron. The ejected electron will have an kinetic energy of the photon energy less the binding energy of the electron.



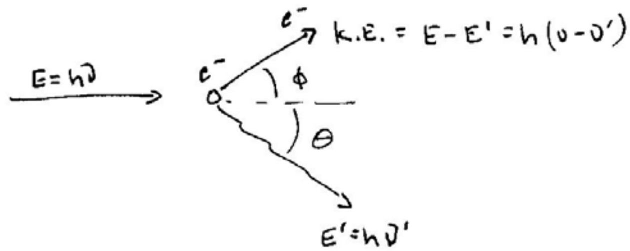
The photoelectric effect increases rapidly with atomic number, Z, and with decreasing energy. The photoelectric effect dominates μ in the lower part of the diagnostic spectrum.



For high Z materials (e.g. Lead, Iodine, Tungsten), the shell energy boundaries are evident in the μ vs. E plots. When the energy gets high enough to make that shell's electrons available to the PE effect (when E exceeds the binding energy), then the probability of a PE interaction increases.



3. Compton Scattering – scatting of photons by an elastic collision with a free electron. Elastic collisions preserve E and momentum (p). For loosely bound electrons or very high energy photons, the equations for free electrons hold reasonably well.



Unknowns: ϕ, θ, E', K.E.

Conservation of energy:

$$\text{K.E.} = E - E' = (m - m_0)c^2$$

where $m = \dfrac{m_0}{\sqrt{1 - v^2 / c^2}}$ is the relativistic mass of the electron

Just a check on this equation … for $v^2 << c^2$, then

$$(m - m_0)c^2 = m(1 - \sqrt{1 - v^2 / c^2})c^2$$

$$\approx m(1 - (1 - \frac{1}{2}\frac{v^2}{c^2}))c^2 = \frac{1}{2}mv^2$$

Conservation of momentum in x and y directions:

$$\frac{E}{c} = \frac{E'}{c}\cos\theta + mv\cos\phi$$

$$\frac{E'}{c}\sin\theta = mv\sin\phi$$

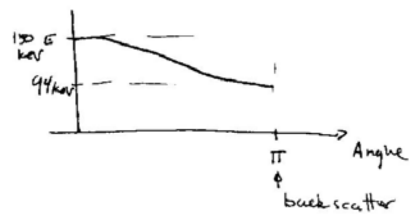
solving these equations we get the energy of the scattered photon:

$$E' = \frac{E}{1 + \dfrac{E}{E_e}(1 - \cos\theta)}$$

where $E_e = m_0c^2 = 511$ keV, the rest energy of an electron.

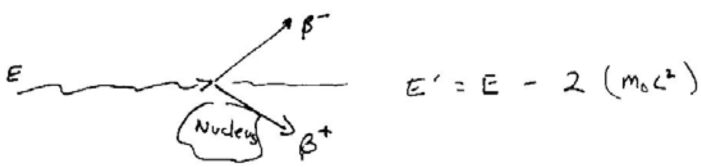
Comments:

- For $E << E_e$, there is very little change in energy with angle.
- For higher E:

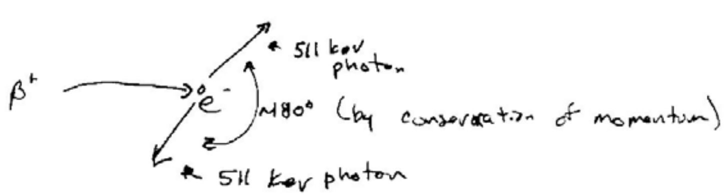


- For low E, scatter is essentially isotropic in angle
- For higher E, scatter is preferentially forward scattered (where there is very little change in photon E).
- It is very hard to discriminate between forward scattered photons and unimpeded photons based on energy.
- μ_{cs} is nearly constant across diagnostic spectrum
- Compton scatter comes mostly from atomic electrons (μ_{cs} is proportional to ρ)
- At higher E, Compton scatter dominates over the PE effect (most important effect in x-ray imaging).

4. Pair Production – the spontaneous creation of an electron/positron pair:



In this interaction, photon energy in transferred to mass energy in the electron and positron. Since the rest energy of each is 511 keV, pair production cannot occur for x-ray photons below 1022 keV (not in the diagnostic spectrum). Positrons will wander around until they bump into an electron, which will result in mutual annihilation and the emission of two 511 keV photons:



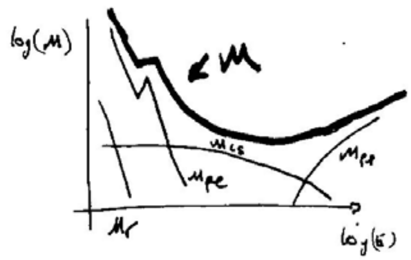
The ejected photons from a positron/electron annihilation is the basis for positron emission tomography [more on this later].

Total Linear attenuation coefficient for photons

Again, the combined coefficient is:

$$\mu(E) = \mu_{rt}(E) + \mu_{pe}(E) + \mu_{cs}(E) + \mu_{pp}(E) + \ldots$$

For example, the combined coefficient for lead is:



An alternate to linear attenuation coefficient is the "mass attenuation coefficient" which is defined as:

$$\tau = \mu / \rho \text{ (units: cm}^2\text{/gm)}$$

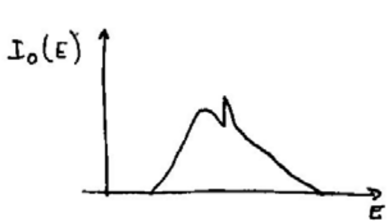
This parameter is convenient when describing the behavior of composite materials with N constituent components:

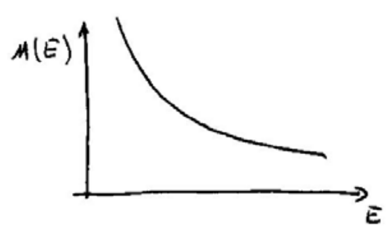
$$\tau = \frac{1}{M}\sum_{i=1}^{N} m_i \tau_i$$

where m_i are the masses of the components and M is the total mass.

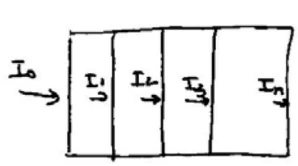
Beam Hardening

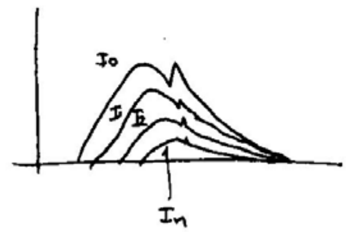
Because the attenuation spectrum is not uniform across the diagnostic energy spectrum, the output spectrum will have a different intensity distribution than the input spectrum, $I_0(E)$.





If we split an object into several smaller parts, and look at then energy spectrum at for each part:





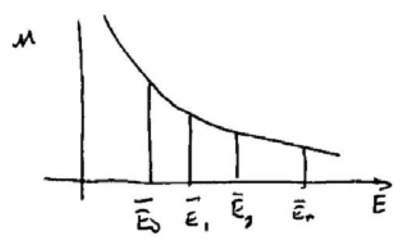
we will find that the mean energy:

$$\overline{E} = \frac{\int EI(E)dE}{\int I(E)dE}$$

will increase (get harder) as we move through the object:

$$\overline{E}_0 < \overline{E}_1 < \overline{E}_2 < ... < \overline{E}_n \, .$$

For medical imaging, this has the unfortunate consequence that a particular tissue type will have a μ that changes as a function of position along the path.



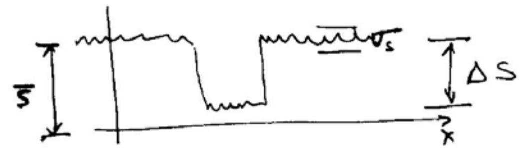
In particular, as we move deeper into the object, we will find that there is less attentuation than expected, given the initial spectrum, $I_0(E)$.

One solution is to make the beam "hard" to begin with. This is often accomplished by filtering out the low E photons with a thin metal plate (often use aluminum).

Compton Scattered X-rays

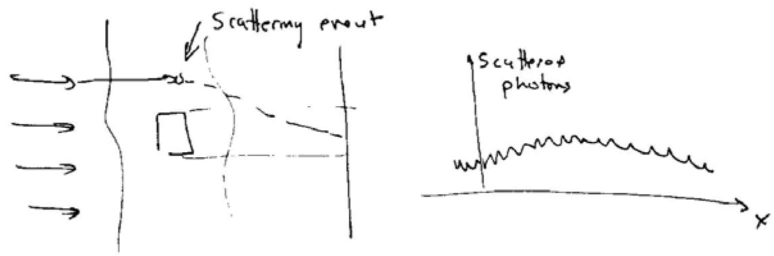
Consider the following object with an x-ray opaque core:

the output image might look like this:



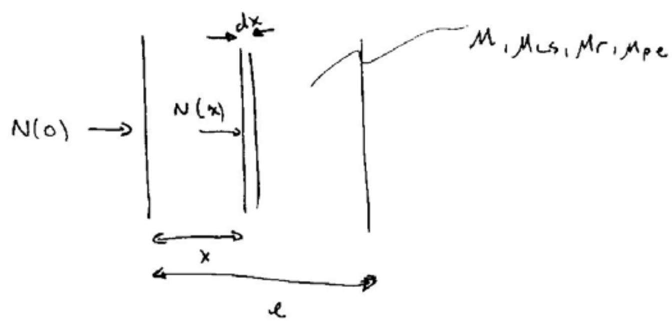
on which we can define a contrast $C = \Delta S / \overline{S}$ and a contrast to noise ratio $CNR = \Delta S / \sigma_s$. Now, consider that the scattered photons – here some fraction of the scattered photons will scatter forward and will generate additional photons in the final image.

The distribution of the scattered photons will look something like the object convolved with the forward scattering distribution. The final image will be the sum of the transmitted photons and the scattered photons.



By increasing $\overline{S}$ and σ_s the scattered photons will reduce both the contrast and the contrast to noise ratio.

How many photons are scattered? (Derived from Macovski, Problem 3.4) Let's look at an object of length *l* having an attenuation coefficient $\mu = \mu_{rt} + \mu_{pe} + \mu_{cs}$. Let *N*(0) be the number photons incident upon the object and that the number of photons that have not interacted at depth *x* is *N*(*x*).



The number of scattered photons in an interval *dx* will be:

$$N_{cs}(x) = \mu_{cs} N(x) dx$$

and the total number of scattered photons will be:

$$N_{cs} = \int_0^l N_{cs}(x) dx$$

$$= \int_0^l \mu_{cs} N(x) dx$$

$$= \int_0^l \mu_{cs} N(0) \exp(-\mu x) dx$$

$$= \frac{\mu_{cs}}{\mu} N(0)(1 - \exp(-\mu l))$$

Note that $N(0)(1\text{-}\exp(-\mu l))$ is the total number of photons that interact with the object.

Noise in X-ray Systems

In an x-ray system, images typically are created from intensity values that are related to the number of photons that strike a detector element in a finite period of time. The photons are generated by electrons randomly striking a source and thus the photons at the detector are also random in nature. We typically describe this kind of random process as one having a rate parameter, λ (units: events/time), and an observation time, T. Let X be the random variable (R.V.) that describes the number of events (photons striking the detector element) in time T.

X will be a Poisson distributed random variable with parameter λT. E.g.

$$X \sim \text{Poisson}(\lambda T)$$

Derivation of Poisson Distribution

Below, we will derive the Poisson distribution from a set of independent Bernoulli R.V.'s. Let Δt be some small time interval and $N = T/\Delta t$ be the number of independent trials. The probability of an event (photon) in interval Δt will be $\lambda\Delta t$. Each Bernoulli trail will then be an R.V.:

$$Y_i \sim \text{Bernoulli}(\lambda\Delta t)$$

$$Y_i = \begin{cases} \text{event A, with probability } p = \lambda\Delta t \\ \text{event B, with probability } q = 1 - p \end{cases}$$

We also assume that Δt is chosen to be small enough so that the probability that there are two events is very small (later we will let Δt go to zero, so this is a non-issue).

Now we consider the sum of the N events, which yields a binomial R.V.

$$X = \sum_{i=1}^{N} Y_i$$

$$X \sim \text{Binomial}(N, \lambda T)$$

The probability density function is $f(x)$ = Probability$\{X = x\}$ (the probability that there were x events in time T). For a binomial R.V., this is derived from the following:

$$\left.\begin{matrix} \text{A} \\ \vdots \\ \text{A} \end{matrix}\right\} \quad x \text{ of event type A}$$

$$\left.\begin{matrix} \text{B} \\ \vdots \\ \text{B} \end{matrix}\right\} \quad N - x \text{ of event type B}$$

which will occur with probability $p^x q^{N-x}$ and there are $\binom{N}{x} = \frac{N!}{x!(N-x)!}$ different ways to get x of event type A. This yields the following p.d.f.:

$$f(x) = \frac{N!}{x!(N-x)!} p^x q^{N-x}$$

Please also observe that

$$\sum_{x=0}^{N} f(x) = 1$$

The mean of X is:

$$\overline{X} = E[X] = \sum_{x=0}^{N} x \frac{N!}{x!(N-x)!} p^x q^{N-x}$$

$$= Np \sum_{x=1}^{N} \frac{(N-1)!}{(x-1)!(N-x)!} p^{x-1} q^{N-x}, \text{ and letting } N' = N, \text{ and } y = x\text{-}1$$

$$= Np \sum_{y=0}^{N'} \frac{N'!}{(y)!(N'-y)!} p^y q^{N'-y} = Np \sum_{y=0}^{N'} f_{N'}(y) = Np$$

In a similar fashion we can show that

$$E\left[X^2\right] = Np + N^2p^2 - Np^2, \quad \text{and}$$

$$\sigma_X^2 = Np(1-p) = Npq$$

Finally, we will let $\Delta t \rightarrow 0$, $N = T/\Delta t \rightarrow \infty$, $p = \lambda\Delta t \rightarrow 0$, and $q \rightarrow 1$. In the following, keep in mind that $q = 1 - p$, $Np = \lambda T$, $N = \lambda T/p$. The Poisson probability distribution is therefore:

$$\lim_{\Delta t \to 0} f(x) = \lim_{\Delta t \to 0} \frac{N!}{x!(N-x)!} p^x q^{N-x}$$

$$= \lim_{\Delta t \to 0} \left[\frac{N(N-1)\cdots(N-x+1)}{N^x}\right]\left[\frac{N^x p^x}{x!}\right]\left[\frac{1}{q^x}\right]\left[q^N\right]$$

$$= \left[\lim_{N \to \infty} \frac{N(N-1)\cdots(N-x+1)}{N^x}\right]\left[\frac{(\lambda T)^x}{x!}\right]\left[\lim_{q \to 1} \frac{1}{q^x}\right]\left[\lim_{p \to 0} (1-p)^{-1/p}\right]^{-\lambda T}$$

$$= [1]\left[\frac{(\lambda T)^x}{x!}\right][1][e]^{-\lambda T}$$

$$= \frac{e^{-\lambda T}(\lambda T)^x}{x!}$$

[The exponential limit comes from $e^{-\varepsilon} \approx 1 - \varepsilon \rightarrow e \approx (1-\varepsilon)^{-1/\varepsilon}$.]

The mean and variance are:

$$\overline{X} = \lim_{\Delta t \to 0} Np = \lim_{\Delta t \to 0} \frac{T}{\Delta t} \lambda \Delta t = \lambda T$$

$$\sigma_X^2 = \lim_{\Delta t \to 0} Npq = \lim_{\Delta t \to 0} \frac{T}{\Delta t} \lambda \Delta t (1 - \lambda \Delta t) = \lambda T$$

Here X is a Poisson R.V. with parameter λT:

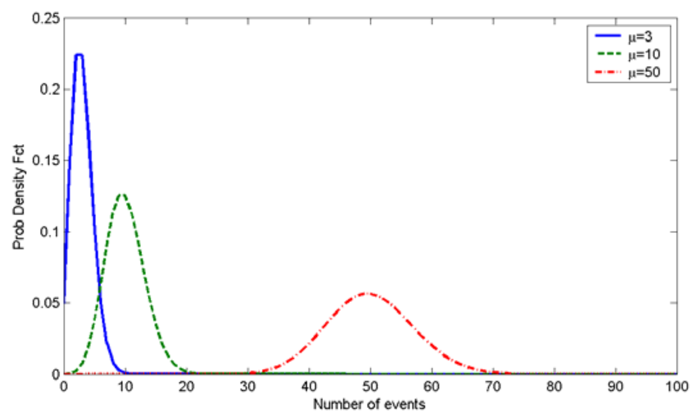
$$X \sim \text{Poisson}(\lambda T).$$

SNR of a Poisson Measurement

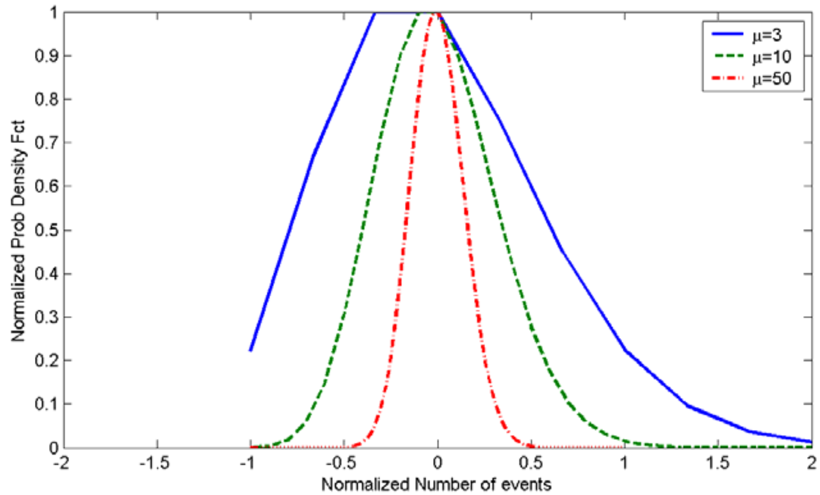
In general, the pixel values in an x-ray image are distributed according to a Poisson R.V. If the mean value of the photon counts for a pixel is μ, then the signal to noise ratio of for that pixel will be:

$$SNR = \frac{\overline{X}}{\sigma_X} = \frac{\mu}{\sqrt{\mu}} = \sqrt{\mu}$$

The SNR increases as the square root of the number of photons. *Thus, the SNR increases as the square root of the dose to the patient.* Finally, by averaging together two neighboring pixels, we can roughly double the photon counts and improve the SNR by $\sqrt{2}$.



The above figure shows Poisson distributions as the mean increases from 3 to 50. We can see that the distribution becomes more symmetric and Gaussian.



The above figure takes Poisson distributions and normalizes them by their mean, that is, we subtract the mean and divide the x-axis by the mean. This plot show demonstrates that the width of the distribution as a fraction of the mean. As the mean gets larger, the distribution gets proportionately narrower – the std. dev. vs. mean ratio is smaller (SNR is higher).

The Relationship of a Poisson Process to the Exponential R.V.

Let *T* be an exponential R.V. that describes the time between events in a Poisson process. The derivation follows. Recall that the probability that an event occurs in interval Δt will be $p = \lambda \Delta t$. Also, note that the probability that no event occurs in interval Δt will be $q = (1 - \lambda \Delta t)$. Now, suppose the we want to know what is the probability that no event occurred between 0 and *t*. This is the same as saying that we have $N = t/\Delta t$ inverval in which no event can occur. If these intervals are independent (that is saying that the photons don't interact with each other or tend to come in groups or whatever), the probability that no event occurred between 0 and *t* will be q^N:

$$\Pr\{\text{no event occurs in } (0,\text{t})\} = (1 - \lambda \Delta t)^{t/\Delta t}$$

we again determine this function as $\Delta t \rightarrow 0$:

$$\Pr\{\text{no event occurs in } (0,\text{t})\} = \lim_{\Delta t \to 0} \left[(1 - \lambda \Delta t)^{-1/(\lambda \Delta t)} \right]^{-\lambda t} = e^{-\lambda t}$$

The probability density function of T, $f(t)$, describes the probability that an event occurs at time t and probability distribution function of T (integral of $f(t)$) describes the probability that an event occurs by time t will be equal to:

$$F(t) = 1 - \Pr\{\text{no event occurs in } (0, \text{t})\} = \begin{cases} 1 - e^{-\lambda t}, \text{ for } t \geq 0 \\ 0, \text{ for } t < 0 \end{cases}$$

and the probability density function is the derivative of this function:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, \text{ for } t \geq 0 \\ 0, \text{ for } t < 0 \end{cases}$$

The exponential R.V. is a continuous R.V. of the times between events and is described as:

$$T \sim \text{Exponential}(\lambda)$$

which has a mean an variance of:

$$\overline{T} = \frac{1}{\lambda}$$

$$\sigma_T^2 = \frac{1}{\lambda^2}$$

Memoryless Property

The exponential R.V. is "memoryless," meaning that distribution (and density) of event times in the future is not affected by past events, that is, at any point in time, the time until the next event is an exponential R.V. with parameter λ. This is the same as saying that just because we haven't seen an event in a long time, we're not more likely to have an event soon. (Just like the "gambler's fallacy.") Specifically,

$$\Pr\{T > t + t_0 \mid T > t_0\} = \Pr\{T > t\}$$

which says "given that an event hasn't occurred by time t_0, the probability that an event will not occur by time ($t_0 + t$) will be same as the probability that no event occurs in $(0,t)$." Proof:

$$\Pr\{T > t + t_0 \mid T > t_0\} = \frac{1 - F(t + t_0)}{1 - F(t_0)}$$

$$= \frac{\exp(-\lambda(t + t_0))}{\exp(-\lambda t_0)} = \exp(-\lambda t) = 1 - F(t)$$

$$= \Pr\{T > t\}$$