Chap. 3 Statistics & Measurement

MICHIGAN STATE UNIVERSITY

Fact: Randomness of decay leads to statistical fluctuations.

Goals:

(I) Characterization of fluctuations in data: The problem is: a set of measurements of one quantity gives a range of answers. How should we describe the *set* ?

Experimental Mean: $\overline{x_e} = \frac{1}{N}\sum_1^N x_i$

Residual: $d_i = x_i - \overline{x_e}$ *note* $\sum_1^N d_i = 0$

Deviation from (*the*) mean $\varepsilon_i = x_i - \overline{x}$
And its variance: $\sigma^2 = \overline{\varepsilon^2} = \frac{1}{N}\sum_1^N \left(x_i - \overline{x}\right)^2$

Sample Variance: $s^2 = \frac{1}{N-1}\sum_1^N \left(x_i - \overline{x_e}\right)^2$ *practical expresssion* $s^2 \approx \overline{x_i^2} - \left(\overline{x_e}\right)^2$

Statistics & Measurement – Models –1–

MICHIGAN STATE UNIVERSITY

(II) Fluctuation Models:

A) The binomial distribution, three parameters in general (a,b,n)

$$(a+b)^n = a^n / 0! + (na^{n-1} b^1)/1! + [n(n-1) a^{n-2} b^2]/ 2! + \ldots [n! a^{n-x} b^x]/ (n-x)!x! + b^n / 0!$$

Special case of a fractional probability, p, where one variable is removed,
$b = p$, $a = 1-p$ so that $a+b = 1$ then the normalized probability function is:

$$(a+b)^n = \Sigma P_n(x) = (1-p)^n + \ldots [n! (1-p)^{n-x} p^x]/ (n-x)!x! \ldots + p^n$$

Thus, if "p" is the probability of success and "n" is the number of tries, then x is the number of successes.

Example 1: Coin flip: $p = ½$ $(1-p) =1/2$

Example 2: Toss die: $p = 1/6$ $(1-p) = 5/6$

Example 3: Probability of radioactive decay:

$$p = (N_0 – N)/N_0 = (1 – e^{-\lambda t}) \quad (1-p) = e^{-\lambda t}$$

Statistics & Measurement – Models –2–

(II) <u>Fluctuation Models</u>:

A) Binomial distribution, special case: n tries, fractional probability p

$$P_n(x) = [n! / (n-x)!x!] \ (1-p)^{n-x} \ p^x$$

$$\text{Mean} = p*n, \qquad \sigma^2 = p\,n\,(1-p), \qquad \sigma^2/\text{mean} = (1-p)$$

Toss die: $p = 1/6$, $(1-p) = 5/6$ …

Toss three dice all at once, chance to get a "5" ?

(the same as three throws of one die to get a "5")

$n=3$, x is number of 5's

$\text{Mean} = p*n = 1/6 * 3 = 1/2 \qquad \sigma^2 = p\,n\,(1-p) = 1/6*3*(5/6) = 5/12$

x = 0	$P_3(0) = [3! / (3-0)!\ 0!]\ (1/6)^0\ (5/6)^3$	0.567
1	$P_3(1) = [3! / (3-1)!\ 1!]\ (1/6)^1\ (5/6)^2$	0.347
2	$P_3(2) = [3! / (3-2)!\ 2!]\ (1/6)^2\ (5/6)^1$	0.0694
3	$P_3(3) = [3! / (3-3)!\ 3!]\ (1/6)^3\ (5/6)^0$	0.00463




$\Sigma P_n(x) = 1$

(II) Fluctuation Models:

B) The Poisson distribution ($p \rightarrow 0$) and large values of n

$P_n(x) = [n!\,(1-p)^{n-x}\, p^x]/(n-x)!x! \rightarrow P_n(x) = [n^x (1-p)^{n-x}\, p^x]/x!$ Since $n-x \sim n$

$P_n(x) = [n^x p^x (1-p)^{n-x}]/x!$

$$n^x p^x = (np)^x = (\bar{x})^x \quad \& \quad p = \bar{x}/n$$

$$P_n(x) = \frac{(\bar{x})^x}{x!}\left(1 - \frac{\bar{x}}{n}\right)^{n-x}$$

Only one parameter, the mean

$$\sum P_P(x) = 1 \qquad \bar{x} = p\,n \qquad \sigma^2 = \bar{x}(1-p) = \bar{x}$$

Asymmetric distribution with a tail on high side

MICHIGAN STATE UNIVERSITY

(II) Fluctuation Models:

B) Example of Poisson Distribution: the birthday problem, how many people in this room have the same birthday?

$$P_P(x) = \frac{\left(\overline{x}\right)^x}{x!} e^{-\overline{x}}$$

$$\overline{x} = pn \qquad p = \frac{1}{365.25}$$

$$P_P(x=0) = \frac{\left(\overline{x}\right)^0}{0!} e^{-\overline{x}}$$

$$P_P(x=1) = \frac{\left(\overline{x}\right)^1}{1!} e^{-\overline{x}}$$

p =	0.002737851
n =	16
<x> = pn =	0.043805613
x	P(x)
0	0.957139995
1	0.041928104
2	0.000918343
3	1.34095E-05
4	1.46853E-07

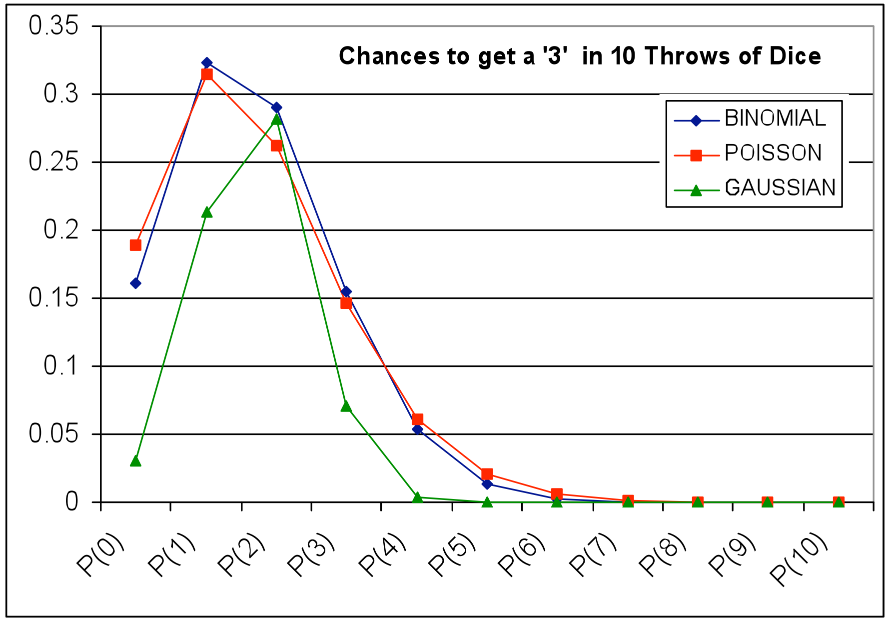
(II) Fluctuation Models:

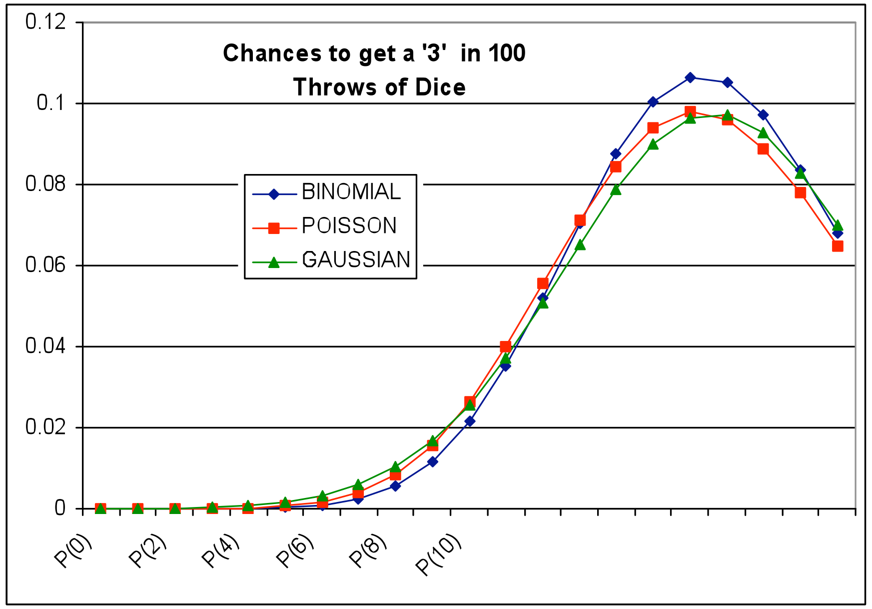
C) The Gaussian approximation, p→0, n>>1 & mean > 20

$$P_G(x) = \frac{1}{\sqrt{2\pi \bar{x}}} e^{-(\bar{x}-x)^2/2\bar{x}} \quad or \quad G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\bar{x}-x)^2/2\sigma^2}$$

Symmetric Continuous distribution (one parameter) also chi-squared distribution:

$$G(\chi) = \frac{2}{\sqrt{2\pi}} e^{-\chi^2/2} \qquad where\ \chi = \left(\bar{x} - x\right)/\sigma$$





Statistics & Measurement –Analysis –1–

(III) Analysis of data

A) Are fluctuations in N-values consistent with statistical accuracy?

Measure: x_e and s^2 Model: mean gives σ^2 … thus the test: Is $s^2 = \sigma^2$?

Define a statistic χ in terms of measurables that will provide a simple test.

$$\chi^2 = \frac{1}{\overline{\overline{x}}_e}\sum_{i=1}^{N}\left(\overline{x}_e - x\right)^2 = \sum_{i=1}^{N}\frac{\left(\overline{x}_e - x\right)^2}{\overline{\overline{x}}_e} = \sum_{i=1}^{N}\frac{\left(\overline{x}_e - x\right)^2}{s^2} = \sum_{i=1}^{N}\left(\frac{\overline{x}_e - x}{s}\right)^2$$

$$\chi^2 = \frac{\sum_{i=1}^{N}\left(\overline{x}_e - x\right)^2}{\overline{\overline{x}}_e} = \frac{s^2(N-1)}{\overline{\overline{x}}_e} \quad \rightarrow \quad \frac{\chi^2}{(N-1)} = \frac{s^2}{\overline{\overline{x}}_e}$$

But in a "good" data set with large N: $s^2 = \sigma^2 = \text{mean}$

$$\therefore \quad \frac{\chi^2}{(N-1)} = 1 \quad or \quad \chi^2 = (N-1)$$

Statistics & Measurement – Analysis –2–

MICHIGAN STATE UNIVERSITY

(III) Analysis of data

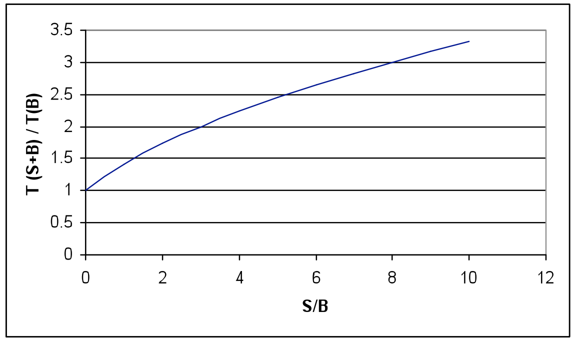
B) Only one measurement, estimate uncertainty?

There is only one measurement, strictly speaking you are out of luck.
However, one can posit that it must be the mean, and σ^2 = mean
One further assumes that the distribution is symmetric: x +/- σ
Fractional error: $\sigma / x = \sqrt{x} / x = 1 / \sqrt{x}$

(IV) Measurements in the presence of Background (itself measureable)
(1) measure background (B); (2) measure source + background (S+B).
S = (S+B) – B … nothing new here but there are two contributions to error.
How should the time be allocated to give minimum uncertainty in S ie s_S ?

Simplest case, constant rates the minimum in s_S comes when:

$$\frac{time_{S+B}}{time_B} = \sqrt{\frac{S+B}{B}} = \sqrt{\frac{S}{B}+1} \quad \& \quad s_S^2 = \frac{B}{T_B}\left(1+\sqrt{\frac{S+B}{B}}\right)$$

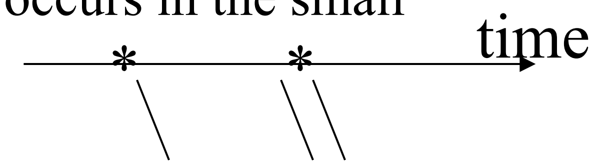


B diminishes in importance for large S

Statistics & Measurement – time distribution

(V) Time between random events, the *Interval distribution*:

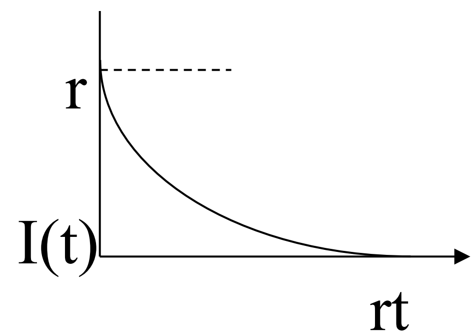
The product of the probability that nothing happens for some length of time, t, times the probability that the next event occurs in the small interval, dt.



$$I_1(t)dt = \left[P_P(x=0)\right]\left[r\ dt\right] \quad where \quad \text{r is average rate}$$

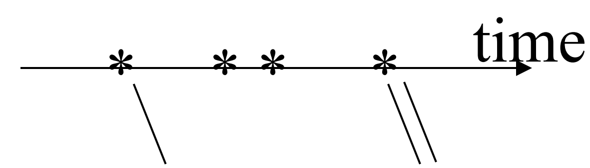
$$I_1(t)dt = \frac{\left(\bar{x}\right)^0}{0!} e^{-\bar{x}}\left[r\ dt\right] \quad where \quad \bar{x} = r * t$$

$$I_1(t)dt = e^{-r\,t}\, r\, dt$$



Similar arguments give the distribution of the j-th event as:

$$I_j(t)dt = \frac{(rt)^{j-1}}{(j-1)!} r\, e^{-r\,t}\ dt$$



which has a peak near the value (rt) = j-1