Poisson distribution

$$P(n|s) = \frac{s^n e^{-s}}{n!}$$

$$P(s|n) = \frac{s^n e^{-s}}{n!}$$ subject to some assumptions

Other probability distributions:

Binomial process.

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

Typical example: selection efficiency

Consider N events n pass a set of selection criteria and $N-n$ fail. What is ε and what is the confidence interval for ε?

$$P(n|N,\varepsilon) = \frac{N!}{n!(N-n)!} \varepsilon^n (1-\varepsilon)^{N-n}.$$

What we want is $p(\varepsilon|n,N)$.

Bayes' Theorem:

$$p(\varepsilon|n,N) = \frac{P(n|\varepsilon,N)\,\pi(\varepsilon)}{\int_0^1 P(n|\varepsilon,N)\,\pi(\varepsilon)\,d\varepsilon}$$

Uniform prior hypothesis: $\pi(\varepsilon) = 1$.

Denominator:

$$\frac{N!}{n!(N-n)!}\int_0^1 \varepsilon^n (1-\varepsilon)^{N-n}\, d\varepsilon$$

Look it up in a table of definite integrals:

$$\int_0^1 x^m (1-x)^p\, dx = \frac{\Gamma(m+1)\Gamma(p+1)}{\Gamma(m+p+2)}$$

$$= \frac{m!\,p!}{(m+p+1)!}$$

So $$\frac{N!}{n!(N-n)!}\int_0^1 \varepsilon^n (1-\varepsilon)^{N-n}\, d\varepsilon = \frac{1}{N+1}$$

$$L = \frac{(N+1)!}{n!(N-n)!}\,\varepsilon^n (1-\varepsilon)^{N-n}$$

Maximize $\log L$ with respect to ε:

$$\log L = \text{const} + n\log\varepsilon + (N-n)\log(1-\varepsilon)$$

$$\frac{\partial}{\partial\varepsilon}\log L = \frac{n}{\varepsilon} - \frac{N-n}{1-\varepsilon} = 0$$

$$\frac{n}{\varepsilon} = \frac{N-n}{1-\varepsilon}$$

$$\frac{\varepsilon}{n} = \frac{1-\varepsilon}{N-n}$$

$$\frac{\varepsilon}{n} + \frac{\varepsilon}{N-n} = \frac{1}{N-n} \quad\Longrightarrow\quad \varepsilon = \frac{n}{N}$$

What if $n = 0$? What is the confidence interval on ε?

$$(N+1)\int_0^{\varepsilon_{up}} (1-\varepsilon)^N d\varepsilon = 1-\alpha$$

Let $u = 1-\varepsilon$

$$(N+1)\int_{1-\varepsilon_{up}}^{1} u^N du = u^{N+1}\Big|_{1-\varepsilon_{up}}^{1}$$

$$= 1 - (1-\varepsilon_{up})^{N+1} = 1-\alpha$$

$$(1-\varepsilon_{up})^{N+1} = \alpha$$

$$1-\varepsilon_{up} = \alpha^{1/N+1}$$

$$\varepsilon_{up} = 1-\alpha^{1/N+1}$$

Example: $N = 10$, $\alpha = 0.1 \Rightarrow$ 90% c.l.

$\varepsilon_{up} = 0.1889$

$P(\varepsilon < 0.1889) = 90\%$.

$N = 100$ $\alpha = 0.1$

$\varepsilon_{up} = 0.0225$

Another example: Rather than measure the angular distribution of say, μ^+, we just count the number in the 'forward direction ($p_z > 0$) and the backward direction ($p_z < 0$).

We observe F in the forward direction and B in the backward direction.

$$N = F + B$$

The expected number of muons observed is Poisson distributed:

$$P(N|\nu) = \frac{e^{-\nu}\nu^N}{N!}$$

Once observed, a muon can only be either in the forward or backward direction

$$P(F|N,f) = \frac{N!}{F!(N-F)!} f^F (1-f)^{N-F}$$

$$= \frac{N!}{F!\,B!} f^F (1-f)^B$$

Probability of observing N, F is

$$P = \frac{e^{-\nu}\nu^N}{N!} \frac{N!}{F!\,B!} f^F (1-f)^B$$

$$= \frac{e^{-\nu f}(\nu f)^F}{F!} \frac{e^{-\nu(1-f)}(\nu(1-f))^B}{B!}$$

Equivalent to two independent poisson processes in each hemisphere.

This can be generalized to describe the contents of a histogram with an arbitrary number of bins, M.

Equivalent descriptions:

1. Poisson distributed N and multinomial distributed $n_1, n_2, \ldots, n_M$
2. Independent Poisson distributions for $n_1, n_2, \ldots, n_M$.

Fitting Histograms.

A histogram takes an interval (y_{low}, y_{high}) and divides it into a number of sub-intervals which do not have to have the same width. We observe the number of events in each bin, $n_1, n_2, \ldots, n_M$.

The expected number of events (non-integer) in each bin is ν_i.

Often we have a model that gives ν_i in terms of some parameters $\vec{\theta}$.

Example: $$\nu_i = \int_{y_i}^{y_{i+1}} f(y;\vec{\theta})\,dy$$

$$\approx f(y_i;\vec{\theta})\,\Delta y$$

Gaussian: $$f(y_i; N, \mu, \sigma) = \frac{N}{\sigma\sqrt{2\pi}} e^{-(y_i-\mu)^2/2\sigma^2}$$

We want to determine N, μ, σ and their confidence intervals.

One approach is to construct a χ^2 statistic and minimize it with respect to N, μ, σ:

$$\chi^2 = \sum \frac{(y_i - \nu_i)^2}{\sigma_i^2}$$

What should we use for σ_i?

When ν_i is large, $\sigma_i \to \sqrt{y_i}$.

$$\chi^2 = \sum \frac{(y_i - \nu_i)^2}{y_i}$$

What if ν_i is not large? We might have $y_i = 0$?

It would be better to use the expected variance instead of arbitrarily skipping bins that contain no entries.

$$\Rightarrow \quad \chi^2 = \sum \frac{(y_i - \nu_i)^2}{\nu_i}$$

But the χ^2 statistic is most meaningful when y_i are gaussian distributed which is not the case for small statistics.

$$L = \prod_i \frac{e^{-\nu_i} \nu_i^{n_i}}{n_i!}$$

$$\log L = \text{const} + \sum \left(-\nu_i + n_i \log \nu_i \right)$$

Example:

$$\frac{\partial}{\partial N}\log L = \sum\left(-\frac{\partial \nu_i}{\partial N} + \frac{n_i}{\nu_i}\frac{\partial \nu_i}{\partial N}\right)$$

$$= \sum\left(\frac{n_i}{\nu_i} - 1\right)\frac{e^{-(y_i-\nu_i)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\Delta y$$

$$= \sum\left(\frac{n_i}{N} - \frac{e^{-(y_i-\nu_i)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\Delta y\right)$$

$$= \frac{\sum n_i}{N} - 1 = 0$$

$$\Rightarrow N = \sum n_i$$

This assumes that essentially all of the distribution is contained within the bins.

$$\sum \frac{e^{-(y_i-\nu_i)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\Delta y = 1 .$$

Unbinned likelihood fit.

We make N measurements y_i but don't assign them to bins. The likelihood function is

$$L = \prod_i \frac{e^{-(y_i-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$

It is important that the likelihood function is normalized with respect to the parameters:

$$\int_{-\infty}^{\infty} L\, d\mu = 1$$

$$\int_0^{\infty} L\, d\sigma = 1$$

$$-\log L = \frac{1}{2}\sum \frac{(y_i-\mu)^2}{\sigma^2} + \frac{N}{2}\log 2\pi + N\log\sigma$$

$$\frac{\partial \log L}{\partial \mu} = \sum \frac{(y_i-\mu)}{\sigma^2} = 0$$

$$\Rightarrow \mu = \frac{\sum \frac{y_i}{\sigma^2}}{\sum 1/\sigma^2} = \frac{\sum y_i}{N}$$

$$\frac{\partial \log L}{\partial \sigma} = \frac{\sum (y_i-\mu)^2}{\sigma^3} - \frac{N}{\sigma} = 0$$

$$\sigma^2 = \frac{\sum (y_i-\mu)^2}{N}$$

Note that this estimate of σ^2 is <u>biased</u>.

$\sigma^2_{corr} = \frac{\sum (y_i-\mu)^2}{N-1}$ is less biased.