Significance of two-parameter coincidence

Lindy Blackburn

September 24, 2015

Consider a background of Poisson-distributed events with a particular rate distribution in threshold parameter ρ, thus the density and cumulative distributions are $d\lambda_{\mathrm{c}}/d\rho_{\mathrm{th}}$ and $\lambda_{\mathrm{c}}(\rho_{\mathrm{th}})$ where we have used subscripts on λ_{c} and ρ_{th} to emphasize the use of cumulate rate (rate of events with $\rho > \rho_{\mathrm{th}}$). We would like to calculate the significance (accidental coincidence probability) of an event from this population falling within T of a time-of-interest t_0. Ordinarily one could pick in advance a single threshold ρ_{th} giving a single rate λ_{c}, then use the Poisson probability of falling within a certain time window $P(\Delta t < T) = 1 - e^{-\lambda_{\mathrm{c}} T} \approx \lambda_{\mathrm{c}} T$ for small P. If many different thresholds are tested, the accidental coincidence probability may be multiplied by a trials factor representing the different effective populations of events.

It is convenient to not have to choose a particular threshold in advance, and thus be able to consider a wide range of possible event rates. In this case, one must generate a single ranking statistic for plausible coincidences between events characterized by the two parameters ρ (or equivalently λ_{c}) and T (closeness to t_0). A natural ranking is the inverse false-alarm probability $R = (\lambda_{\mathrm{c}} T)^{-1}$. $\lambda_{\mathrm{c}} T$ was the original accidental coincidence probability P for a single threshold, but in this case we must add up contributions to accidental coincience from all possible combinations of λ_{c} and T in order to get a faithful representation of the probability of a coincidence happening with greater R than our event under consideration. We can calculate the expected number of more highly-ranked events,

$$N(R > \lambda_{\mathrm{c}} T) = \int_0^{\infty} d\lambda \int_0^{\lambda_{\mathrm{c}} T/\lambda} dt\, e^{-t\, d\lambda} \tag{1}$$

By representing the calculation as a sum over slivers of $d\lambda$, we can conveniently bypass details about the actual shape of $\lambda_{\mathrm{c}}(\rho_{\mathrm{th}})$. Each sliver actually has the same Poisson distribution $d\lambda\, e^{-t\, d\lambda} = d\lambda + O(d\lambda^2)$ since they all cover the same amount of differential rate. However the rank itself is determined by the cumulative rate, which sets the limit of integration. The exponential reduces to first-order in infinitesimal $d\lambda$ (flat) and the integral becomes,

$$N(R > \lambda_{\mathrm{c}} T) = \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda\, \frac{\lambda_{\mathrm{c}} T}{\lambda} = \lambda_{\mathrm{c}} T \ln\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right) \tag{2}$$

Where $\lambda_{\max}$ and $\lambda_{\min}$ are necessary for convergence.

$\lambda_{\max}$ is naturally constrained by the production threshold of the events, or by the minimum measurable coincidence time $\lambda_{\max} T_{\min} = \lambda_{\mathrm{c}} T$. We can also choose a maximum coincidence window $T_{\max}$, up to the live-time of the experiment, to set $\lambda_{\min} T_{\max} = \lambda_{\mathrm{c}} T$. Events from $0 < \lambda < \lambda_{\min}$ will still contribute to the accidental coincidence probability but subject to a bounded interval of time $T_{\max}$. Therefore we need to add a constant to the expectation value equal to $\lambda_{\mathrm{c}} T$. Under these constraints the expected number becomes,

$$N(R > \lambda_{\mathrm{c}} T) = \lambda_{\mathrm{c}} T \left[1 + \ln\left(\frac{\lambda_{\max} T_{\max}}{\lambda_{\mathrm{c}} T}\right)\right] \quad \text{or} \tag{3}$$

$$N(R > \lambda_{\mathrm{c}} T) = \lambda_{\mathrm{c}} T \left[1 + \ln\left(\frac{T_{\max}}{T_{\min}}\right)\right] \tag{4}$$

depending on choice of using $\lambda_{\max}$ or $T_{\min}$. A two-sided coincidence window will multiply N by a trials factor of two. The accidental coincidence probability $P \approx N$ for small N.