Calculating Efficiency Uncertainties

Louise Heelan

August 14, 2009

Efficiency Errors

References

Poisson Statistics

Binomial Statistics

Bayesian Statistics

Results

Summary

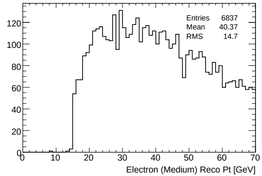
Efficiency Errors

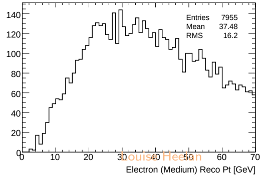
o previously presented trigger efficiency curves (vs pt, eta, phi) for top trigger selections with default ROOT error bars

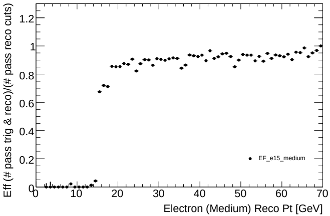
o got me thinking about how to correctly calculate the statistical errors for bin-by-bin efficiency and overall efficiency

o not as trivial as first expected ...

o lots of papers about this topic, so here is a summary of what I found ...







References

G. Cowan, *Error Analysis for efficiency*, July 28, 2008 (ATLASStatisticsFAQ wiki)

D. Casadei, *How to measure efficiency*, July 12, 2009 (ATLASStatisticsFAQ wiki)

T. Ullrich and Z. Xu, *Treatment of Errors in Efficiency Caculations*, January 17, 2007 (physics/0701199v1)

Notation

- let k (k_i) = total number of selected events (in bin i)
- let n (n_i) = total number of events (in bin i)
- want to measure efficiency $\epsilon = k/n$ (efficiency for bin i, $\epsilon_i = k_i/n_i$)
- and assign an uncertainty to the calculated ϵ (based on the measured quantities)

From here on just use the notation k, n, but can apply to individual bins k_i, n_i

Default in ROOT

- TH1::Divide supposed to be standard propagation of errors assuming independent variables m, n (which is not true as these are highly correlated, m and subset of n)
- but when I checked it was just the square-root of the bin efficiency (very wrong, ex. 0.8 ± 0.89) ??

Poisson Statistics

- for large sample limit uncertainties are:

$$\begin{aligned} k \pm \sigma_k &= k \pm \sqrt{k} \\ n \pm \sigma_n &= n \pm \sqrt{n} \end{aligned}$$

- assuming k and n independent (incorrect, k a subset of n):
  - estimator of (unknown) efficiency : $E(\epsilon) = \hat{\epsilon} = k/n$
  - variance : $V(\hat{\epsilon}) = \sigma_{\hat{\epsilon}}^2 = \hat{\epsilon}^2(\frac{1}{k} + \frac{1}{n})$
- problems with model:
  - n fixed quantity, well defined and known ($\sigma_n = 0$), so $V(\hat{\epsilon}) = k/n^2$
  - n and k highly correlated
  - limiting cases where $k = 0 \rightarrow \hat{\epsilon} \pm 0$, even if observed one event ($n = 1$) and it fails ($k = 0$) we know with complete certainty (zero error) the efficiency is zero

Binomial Statistics

o k successes out of n independent trials, k binomially distributed

o probability of success is a function of *true* efficiency ϵ :

$$P(k; n, \epsilon) = \frac{n!}{k!(n-k)!}\epsilon^k(1-\epsilon)^{n-k}$$

o for binomially distributed k:

- estimator : $k = n\epsilon$
- variance : $V(k) = \sigma_k^2 = n\epsilon(1-\epsilon)$

o for unknown quantity ϵ (assuming n known with zero error, has no dependence on parameter of interest ϵ):

- assume usual estimator : $E(\epsilon) = \hat{\epsilon} = k/n$
- variance via propagation of errors : $V(\hat{\epsilon}) = \sigma_{\hat{\epsilon}}^2 = V(k/n) = \frac{V(k)}{n^2} = \frac{\epsilon(1-\epsilon)}{n}$

o in ROOT apparently can use TH1:Divide option "B" for errors to be calculated using this binomial method (I tried it ... and it didn't work ...)

Binomial Statistics

- problems with model:
  - ϵ is the quantity we are determining through measurements of k and n, yet uncertainty of measured ϵ depends on true ϵ
  - limiting cases where $k = 0 \rightarrow \hat{\epsilon} \pm 0$, as with Poisson model
  - if $k = n \rightarrow \epsilon = 1$ error is zero

Bayesian Statistics

o determine probability that $\hat{\epsilon}$ is the true efficiency given the measurements of k and n:

$$P(\epsilon; k, n) = \frac{P(k; \epsilon, n) P(\epsilon; n)}{C}$$

o where $P(k; \epsilon, n)$ is the probability of k given a certain ϵ and n, and is the binomial probability given two slides ago

o C is the overall normalization constant (proven in Ullrich and Xu) $C = \frac{1}{n+1}$

o and $P(\epsilon; n)$ is the probability of ϵ given a value of n, prior before any measurements taken leads us to assume $P(\epsilon; n)$ is uniform [0, 1] (ϵ can take any value between 0 and 1 regardless of n)

o the final efficiency probability function is:

$$P(\epsilon; k, n) = \frac{(n+1)!}{k!(n-k)!} \epsilon^k (1-\epsilon)^{n-k}$$

Bayesian Statistics: $n = 10$

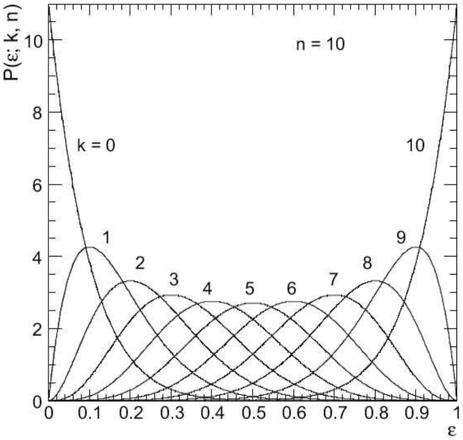


Figure 1: The probability density function $P(\varepsilon; k, n)$ for $n = 10$ and $k = 0, 1, \ldots, 10$.

Bayesian Statistics: n = 100

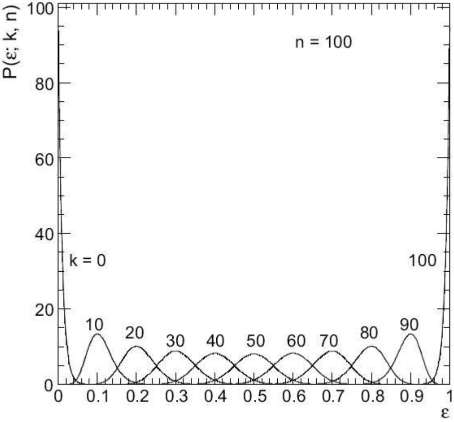


Figure 2: The probability density function $P(\varepsilon; k, n)$ for $n = 100$ and $k = 0, 10, \ldots, 100$.

Bayesian Statistics: Efficiency

- for the efficiency we can calculate the moments, and one find:
  - mean: $\bar{\epsilon} = \frac{k+1}{n+2}$
  - mode (most probably value, $dP/d\epsilon = 0$) : $\text{mode}(\epsilon) = \frac{k}{n}$

- interesting to note:
  - mean value not as typically expected, mode is what we call efficiency, and mode $=$ mean when $n = 2k$ (mid-value, symmetric distribution), and for large n
  - papers claim mean is biased for small n
  - for small n and/or small k, the values for the mode and mean are different

Bayesian Statistics: Variance

- the variance of the efficiency is (proof in Ullrich and Xu):

$$V(\epsilon) = \frac{(k+1)(k+2)}{(n+2)(n+3)} - \frac{(k+1)^2}{(n+2)^2}$$

- interesting points about the variance:
  - now note that variance behaves correctly for the extreme cases of $k = 0$ and $k = n$
  - for $k = n$ as n gets large $V(\epsilon) \rightarrow 1/n^2$
  - for $n = 0$ the mean and variance is non-zero ($\bar{\epsilon} = 1/2$ and $V(\epsilon) = 1/12$, mean and variance for a uniform distribution with range (0, 1)).
  - dependent only on variables that are measured or known
  - does not work if $k > n$ (variance negative)

according to references listed this is the correct treatment of the statistical uncertainty of the efficiency

Bayesian Statistics: Mode Vs Mean

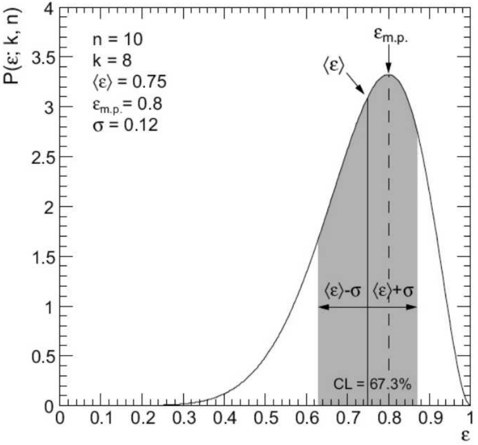


Figure 3: Efficiency probability density function $P(\varepsilon; 8, 10)$. The solid vertical li
depicts the mean value, the dashed line the most probable value. The gray shad
region corresponds to plus/minus one standard deviation (Eq. 19) around the mea

Results

o example of different techniques for my top trigger analysis (105200, r635, ≈30k events)

o two examples of a particular bin (low and high statistics)

Method	Numerator	Denominator	Mean (Mode)	Variance	Uncertainty σ
Poisson	1	45	0.0222	0.00050	0.02246
Binomial	1	45	0.0222	0.00048	0.02197
Bayesian	1	45	0.04255 (0.0222)	0.00085	0.02913

Method	Numerator	Denominator	Mean (Mode)	Variance	Uncertainty σ
Poisson	100	106	0.9433	0.01729	0.13151
Binomial	100	106	0.9433	0.00050	0.02244
Bayesian	100	106	0.9352 (0.9433)	0.00056	0.02358

Summary

o still not sure if Bayesian method is right - which value of efficiency to use? mode or mean, these vary greatly at low n and k

o statistics experts seem to agree Bayesian method is the correct way of calculating the variance on the efficiency

o does this properly take into account correlations between k and n?

o interesting to think about ...