CERN - Data Handling Division
DD/80/1
January 1980

ERRORS ON RATIOS OF SMALL NUMBERS OF EVENTS

Frederick James
Data Handling Division
CERN, Geneva

and

Matts Roos
Department of High Energy Physics
University of Helsinki, SF-00170 Helsinki 17

DD-ag

ABSTRACT

The correct expression for the error on the ratio of numbers of events is given. Values are tabulated for small numbers of events. Conventionally the normal approximation is used, yielding too small errors. We re-evaluate the errors on some published experimental ratios in e^+e^- physics and in neutrino physics.

When the result of the measurement of a physical quantity is published as $R=R_0 \pm \sigma_0$ without further explanation, it is implied that R is a Gaussian-distributed measurement with mean R_0 and variance σ_0^2 . This allows one to calculate various confidence intervals of given "probability", i.e. the "probability" P that the true value of R is within a given interval. P is given by the area under the corresponding part of the Gaussian curve, and is the basis of well-known rules-of-thumb such as "the probability of exceeding two standard deviations is 5%".

In many cases, the measurement R is not Gaussian-distributed, but its variance σ^2 can still be determined. This is true in particular for Poisson-distributed measurement (e.g. number of events counted in a given time period) where the variance is equal to the number of events counted, even though the distribution is only approximately Gaussian for large numbers of events. In this case it is conventional to quote the error as the square root of the variance, since this is indeed the root-mean-square error. However, the extension to confidence intervals should now be made using the Poisson distribution rather than the Gaussian.

When the quantity R is obtained as the ratio of two numbers $R=N \pm \sigma_N / D \pm \sigma_D$, extreme care must be taken. In particular, if N and D are either Gaussian- or Poisson-distributed, then not only is R not Gaussian- (or Poisson-) distributed, but its variance (and therefore the r.m.s. error in R) is always infinite, essentially because of the finite probability at $D=0$. Nevertheless, confidence intervals can be determined, either exactly, using a simple technique described below, or approximately, using the customary formula :

$$\Delta_R = \frac{N}{D} \sqrt{\left(\frac{\sigma_N}{N}\right)^2 + \left(\frac{\sigma_D}{D}\right)^2} . \quad (1)$$

$R \pm \Delta_R$ is then said to be a 68.3% confidence interval. It is the purpose of this article to point out that the above approximation may be very poor for small numbers of events (or large relative errors $\frac{\sigma_N}{N}$ and $\frac{\sigma_D}{D}$). Indeed, in view of the fact that the true r.m.s. error in R is infinite even for small errors σ_N and σ_D , one might wonder why the above approximation ever works. To see this, let us assume that N and D are Gaussian-distributed and consider the distribution of R . Then if the measured values N_0 and D_0 of N and D are large, the distribu-

tion of R near the value $R_0 = N_0/D_0$ appears to be of Gaussian shape, with standard deviation given by the traditional formula (1). Far away from this value, R deviates more and more strongly from a Gaussian distribution. However, for large N_0 and D_0 , the probability content in the tails is small, and one- or two-standard-deviation confidence intervals based on the approximate Gaussian behaviour of R are then acceptable.

The motivation of this article is that in present-day high energy physics, ratios are often formed using very small numbers of events. In that case the approximation (1) is not justified, the problem must be treated in a totally different way, and as we shall see, the 68.3% confidence region $(R - \Delta_R, R + \Delta_R)$ is strongly underestimated by the Δ_R expression in Eq. (1).

To derive the correct expression we note that only two kinds of events enter into consideration : N-events and D-events, and since we lose no generality by considering the sum $N+D$ as fixed, the problem is binomial. Let the probability of observing one N-event be p , then the observation of N N-events and D D-events sets confidence limits on p . The Clopper-Pearson confidence limits (C_-, C_+) for p at confidence level $\beta = (1-\alpha)$ are given by [2] :

$$C_- (\alpha, N, N+D) = \frac{N}{N+(D+1) f_{1-\frac{\alpha}{2}, 2(D+1), 2N}} \quad (2)$$

$$C_+ (\alpha, N, N+D) = 1 - C_- (\alpha, D, N+D),$$

where $f_{1-\frac{\alpha}{2}, n, m}$ is the $(1-\frac{\alpha}{2})$ -quantile of Fishers F-distribution with degree of freedom (n, m) .

From this one finds that the β confidence interval for R is simply

$$\left(\frac{C_-}{1-C_-}, \frac{C_+}{1-C_+} \right) \quad (3)$$

In Table I we give these confidence intervals for $\beta=68.3\%$ and $\beta=90\%$ for a few small values of N and D .

Although the above results enable us to calculate exact confidence intervals for any number of observed events, they are valid only for the ideal case when no background events are present. We give here an approximate procedure for taking account of background events when the expected number of such events is known.

With background present, we are in fact measuring

$$R' = \frac{N}{D} = \frac{S_N + B_N}{S_D + B_D}$$

where S is the number of events in the signal, and B is background. We wish to establish confidence limits on

$$R = \frac{S_N}{S_D}.$$

Using, as before, the fact that the total number of events is fixed, $S_N + B_N + S_D + B_D = k$, we obtain :

$$R = \frac{R' - \frac{B_N}{k} (1+R')}{1 - \frac{B_D}{k} (1+R')} \quad (4)$$

As this is a general relation between R and R', it holds also for $R_{\pm}$ and $R'_{\pm}$. Our procedure is therefore to determine first $R'_{\pm}$ using the technique described previously with the measured values of N and D, then use the above formula to determine $R_{\pm}$. Unfortunately B_N and B_D are not known, so we must use here the known expectations $\hat{B}_N$ and $\hat{B}_D$ (the calculated backgrounds). This approximation is an optimistic one, since we neglect fluctuations in the background about its expected value. Such a procedure must therefore yield confidence intervals smaller than the exact confidence intervals, but approaching the exact intervals in the limit of very small background.

To illustrate the importance of the above remarks, we re-evaluate some published ratios below.

Brandelik et al. [3] test the hypothesis that the weak current for events of type $e^+e^- \rightarrow e^+K^{\pm} + \text{no charged} + 0-2 \text{ gammas}$ are due to the conventional Cabibbo

current and not to the GIM mechanism. Forming the ratio

$$R_{K\pi} = \frac{N_{eK^{\pm}X}}{D_{e\pi^{\pm}X}} \cdot C ,$$

where C is a correction factor for decay losses and background, they quote

$$R_{K\pi}(E_{cm}=3.99-4.52 \text{ GeV}) = \frac{1}{24} \cdot 1.44 = 0.06 \pm 0.06 ,$$

$$R_{K\pi}(E_{cm}=4.52-5.2 \text{ GeV}) = \frac{1}{16} \cdot 2.24 = 0.14 \pm 0.14 .$$

From the expressions (2) and (3) the correct 68.3% confidence intervals (assuming C to be known without errors) should be $0.06^{+0.15}_{-0.05}$ and $0.14^{+0.36}_{-0.12}$, respectively.

Imlay et al. [4] study the neutrino- produced prompt muon pairs and form the ratios $R_L = N(\mu^+ \mu^-) / D(\mu^+ \mu^-)$ for three different absorption lengths $L=31, 61$ and 120 cm at $p_\mu > 10$ GeV. In this case the observed number of events have to be corrected for background. The experimental figures are :

$L = 31$	$N = 8,$	$\hat{B}_N = 1.6,$	$D = 50,$	$\hat{B}_D = 3.5$
$L = 61$	$N = 4,$	$\hat{B}_N = 1.5,$	$D = 23,$	$\hat{B}_D = 2.6$
$L = 120$	$N = 6,$	$\hat{B}_N = 3.2,$	$D = 32,$	$\hat{B}_D = 7.7$

The normal approximation (1) gives [4] :

$$\begin{aligned} R_{31} &= (8-1.6)/(50-3.5) = 0.138 \pm 0.065 \\ R_{61} &= (4-1.5)/(23-2.6) = 0.123 \pm 0.100 \\ R_{120} &= (6-3.2)/(32-7.7) = 0.115 \pm 0.110, \end{aligned}$$

where the errors are taken to be 68.3% confidence intervals.

From this Imlay et al. conclude that R_L is L-independent. Using equations (2, 3, 4) one obtains :

$$\begin{aligned} R_{31} &= 0.138^{+0.112}_{-0.061} \\ R_{61} &= 0.123^{+0.174}_{-0.095} \\ R_{120} &= 0.115^{+0.165}_{-0.096} \end{aligned}$$

Thus the Imlay et al. conclusion is still consistent with the data, but it is now weakened by the larger errors (which are still underestimates of the exact errors).

Sinclair [5] tests the GIM prediction for the ratio of neutrino-produced events with neutral strange particles to μ^+e^- events. The observation is quoted as

$$\frac{18}{12} = 1.5 \pm 0.5$$

consistent with the GIM prediction 2. However, the statistically correct result is $1.5^{+0.8}_{-0.5}$, thus even closer to the GIM prediction.

The same experiment sees what is interpreted as either N=1 candidate for a μ^+e^+ event over a background of $\hat{B}_N=0.4$, or N=2 candidates over a background of $\hat{B}_N=1.3$. The result quoted is, using D=12 μ^+e^- events ($B_D=0$)

$$\frac{\sigma(\mu^+e^+)}{\sigma(\mu^+e^-)} \lesssim 0.2, \quad 90\% \text{ C.L.}$$

Using equations (2, 3, 4) we find the one-sided 90% confidence limits are at least 0.42 and 0.48 for N=1 and 2, respectively.

R. Brandelik et al. [6] find 30 events of the reaction e^+e^-

$$e^+e^- \rightarrow J/\psi(3100) \rightarrow f\gamma$$

$$\quad \quad \quad \hookrightarrow \pi^+\pi^-$$

and at most 4 events of the reaction

$$e^+e^- \rightarrow J/\psi(3100) \rightarrow f'\gamma$$

$$\quad \quad \quad \hookrightarrow K^+K^-$$

They conclude that

$$R = \frac{\Gamma(J/\psi \rightarrow f'\gamma)}{\Gamma(J/\psi \rightarrow f\gamma)} \leq A \cdot \frac{4}{30} \approx 0.3,$$

where $A \sim 2.25$. In fact, the upper 90% confidence limit from equations (2) and (3) is 0.64.

It would be easy to continue this list with other τ studies, neutrino studies, branching ratios of the F or the $J/\psi(3100)$, etc. However, we think our point has been sufficiently well illustrated :

- it is not safe to use the normal approximation for ratios of small numbers of events;
- the statistically correct confidence limits can be taken from Table I or evaluated from Eqs. (2) and (3);
- when background has to be subtracted, Eq. (4) together with Table I or Eqs. (2) and (3) gives inner bounds to the confidence limits.

REFERENCES

- [1] W.T. Eadie, D. Drijard, F.E. James, M. Roos and B. Sadoulet, 'Statistical Methods in Experimental Physics', North-Holland Publ. Co., Amsterdam (1971).
- [2] C.J. Clopper and E.S. Pearson, "Biometrika 26", 404 (1934);
M. Hollander and D.A. Wolfe, 'Nonparametric Statistical Methods', J. Wiley, New-York (1973). These limits are tabulated in R.E. Odeh, D.B. Owen, Z.W. Birnbaum, L. Fisher, 'Pocket Book of Statistical Tables', Marcel Dekker Inc., New-York and Basel (1977).
- [3] R. Brandelik et al., Phys. Letters 70B, 125 (1977).
- [4] R. Imlay et al., Neutrino 78 Proc., Purdue University (1978), p. 507.
- [5] D. Sinclair, Neutrino 78 Proc., Purdue University (1978), p. 543.
- [6] R. Brandelik et al., Phys. Letters 74B, 292 (1978).

Table I

Upper β -confidence limits (two-sided) for the ratio $R=N/D$.

$\beta = 68.3\%$

(Elements above the diagonal are of course reciprocals of the corresponding elements below the diagonal)

D \ N	0	1	2	3	4	5	6	7	8	9	10
1	5.309 0	11.097 .090									
2	1.512 0	2.968 .059	4.400 .227								
3	0.848 0	1.622 .044	2.373 .172	3.115 .321							
4	0.585 0	1.103 .035	1.601 .138	2.091 .259	2.578 .388						
5	0.445 0	.832 .029	1.202 .115	1.564 .218	1.924 .326	2.282 .438					
6	0.359 0	.667 .025	.960 .099	1.246 .188	1.530 .282	1.811 .379	2.092 .478				
7	0.301 0	.556 .022	.798 .087	1.034 .165	1.267 .248	1.499 .334	1.729 .422	1.959 .510			
8	0.259 0	.477 .019	.682 .077	.883 .147	1.081 .222	1.277 .299	1.472 .377	1.667 .457	1.861 .537		
9	0.227 0	.417 .017	.596 .070	.770 .133	.942 .200	1.112 .270	1.281 .341	1.449 .414	1.617 .487	1.784 .560	
10	0.202 0	.370 .016	.529 .064	.682 .121	.834 .183	.984 .247	1.133 .312	1.281 .378	1.429 .445	1.576 .512	1.723 .580

$\beta = 90\%$

D \ N	0	1	2	3	4	5	6	7	8	9	10
1	19.000 0	38.493 .026									
2	3.472 0	6.388 .017	9.245 .108								
3	1.714 0	3.022 .013	4.284 .083	5.529 .181							
4	1.115 0	1.919 .010	2.685 .067	3.438 .148	4.184 .239						
5	0.821 0	1.391 .009	1.930 .056	2.457 .125	2.978 .203	3.496 .286					
6	0.647 0	1.086 .007	1.498 .049	1.889 .108	2.294 .177	2.687 .249	3.077 .325				
7	0.534 0	.889 .006	1.220 .043	1.542 .096	1.859 .156	2.172 .221	2.484 .289	2.794 .358			
8	0.454 0	.752 .006	1.028 .038	1.296 .086	1.558 .140	1.818 .199	2.077 .260	2.333 .322	2.589 .386		
9	0.395 0	.651 .005	.887 .034	1.116 .077	1.340 .127	1.561 .180	1.781 .236	2.000 .293	2.217 .352	2.434 .411	
10	0.350 0	.573 .005	.780 .031	.979 .071	1.174 .116	1.367 .165	1.557 .216	1.747 .269	1.936 .323	2.124 .377	2.312 .433