LA-UR-04-6722

*Approved for public release;
distribution is unlimited.*

Title:	Probabilistic interpretation of Peelle's pertinent puzzle and its resolution
Author(s):	Kenneth M. Hanson Toshihiko Kawano Patrick Talou
Submitted to:	Presented at the International Conference on Nuclear Data for Science and Technology (ND2004), Santa Fe, September 26-October 1, 2004

Los Alamos
NATIONAL LABORATORY

Probabilistic interpretation of Peelle's pertinent puzzle and its resolution

Kenneth M. Hanson,[1] Toshihiko Kawano, and Patrick Talou

Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract. Peelle's Pertinent Puzzle (PPP) states a seemingly plausible set of measurements with their covariance matrix, which produce an implausible answer. To answer the PPP question, we describe a reasonable experimental situation that is consistent with the PPP solution. The confusion surrounding the PPP arises in part because of its imprecise statement, which permits to a variety of interpretations and resulting answers, some of which seem implausible. We emphasize the importance of basing the analysis on an unambiguous probabilistic model that reflects the experimental situation. We present several different models of how the measurements quoted in the PPP problem could be obtained, and interpret their solution in terms of a detailed probabilistic analysis. We suggest a probabilistic approach to handling uncertainties about which model to use.

PEELLE'S PERTINENT PUZZLE

Peelle's original statement [1] of the PPP is as follows:

"Suppose we are required to obtain the weighted average of two experimental results for the same quantity. The first result is 1.5, and the second result is 1.0. The full covariance matrix of these data is believed to be the sum of three components. The first component is fully correlated with standard error 20% of each representative value. The second and third components are independent of the first and each other, and correspond to 10% random uncertainties in each result.

The weighted average from the least-squares method is 0.88 ± 0.22, a value outside the range of the input values! Under what conditions is this the reasonable result that we sought to achieve by use of an advanced data reduction technique?"

The PPP effect has been observed in numerous evaluations of nuclear cross sections. [2, 3, 4, 5, 6]

Initially, the answer stated in the PPP seems implausible. The seeming paradox arises because the statement of the puzzle is ambiguous. For example, it is not stated whether the correlated uncertainty contributes to the measurements in an additive or multiplicative manner. We propose a plausible experimental situation that would correctly yield the answer of 0.88. Alternative interpretations of the PPP may be more appropriate in the context of nuclear cross-section evaluation. A precisely-stated uncertainty model clarifies the probabilistic approach that needs to be taken.

---

[1] Corresponding e-mail address: kmh@lanl.gov   LA-UR-04-6722

Standard solution

We designate the two quantities that are measured by x_1 and x_2, and represent them by the vector $\mathbf{x} = [x_1 \ x_2]^T$. The measurements $m_1 = 1.5$ and $m_2 = 1.0$ are represented by $\mathbf{m} = [m_1 \ m_2]^T$. The measurements have relative independent standard errors of $\rho_1 = \rho_2 = 0.1$, and a relative common error $\rho_c = 0.2$. Assuming that normal distributions are appropriate, the probability density function (pdf) for $\mathbf{x}$ is given by [7, 8]

$$p(\mathbf{x}|\mathbf{m}) \propto \exp\left\{-\tfrac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})\right\}, \quad (1)$$

where the covariance matrix is

$$\mathbf{C} = \begin{pmatrix} m_1^2(\rho_1^2+\rho_c^2) & m_1 m_2 \rho_c^2 \\ m_1 m_2 \rho_c^2 & m_2^2(\rho_1^2+\rho_c^2) \end{pmatrix}. \quad (2)$$

The argument of the exponential in (1) is $\tfrac{1}{2}\chi^2$, generalized to include correlations. Figure 1a shows the joint distribution for x_1 and x_2, considered as separate variables. The strong correlation between x_1 and x_2 is evident from the tilt of the contours.
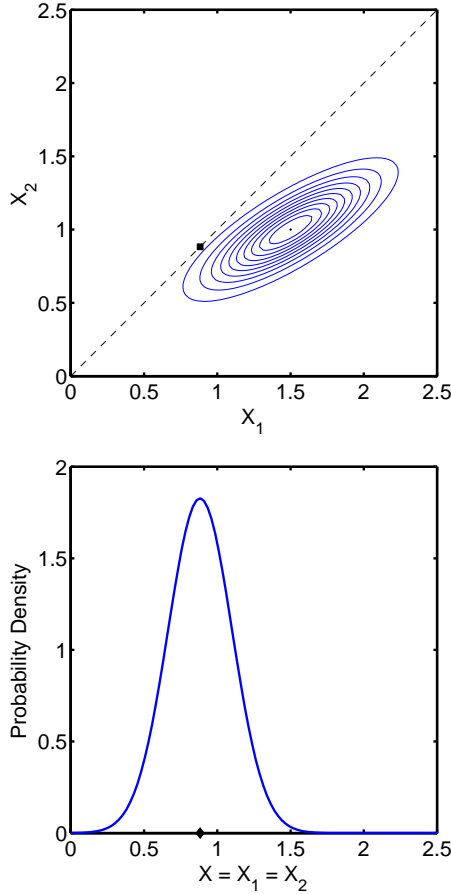
Because x_1 and x_2 are the same quantity, we set $x = x_1 = x_2$. Figure 1b shows the resulting one-dimensional distribution $p(x)$, which is a normal distribution centered on $x = 0.882$ and with a standard deviation of 0.228.

An identical result is given by the familiar least-squares solution, which maximizes the probability (1):

$$\mathbf{x} = (\mathbf{G}^T\mathbf{C}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{C}^{-1}\mathbf{m}, \quad (3)$$

and with the covariance $\mathbf{V} = (\mathbf{G}^T\mathbf{C}^{-1}\mathbf{G})^{-1}$, where $\mathbf{G}$ is the sensitivity matrix, i.e, the matrix of derivatives of measured variables $\mathbf{m}$ with respect to the inferred variables $\mathbf{x}$; $\mathbf{G} = [1 \ 1]$.
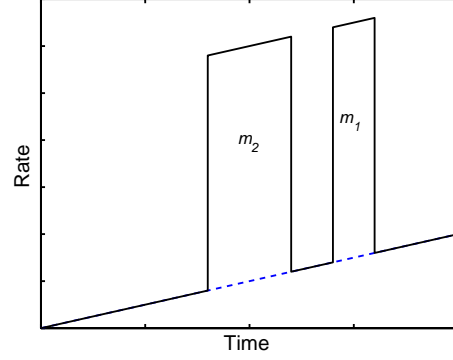
FIGURE 1. (a, top) Contour plot of the joint distribution $p(x_1, x_2 | \mathbf{m})$ given by Eq. (1) and (b, bottom) the distribution along the diagonal for $p(x = x_1 = x_2 | \mathbf{m})$, centered at 0.882.

FIGURE 2. Schematic view of an experimental situation that could yield data consistent with the assumptions made in the standard PPP solution given by Eqs. (1) and (2).

Referring to the logarithm of $p(x, \Delta)$ as $-\varphi$,

$$2\varphi = \frac{(x - m_1 - \Delta)^2}{\sigma_1^2} + \frac{(x - m_2 - \Delta)^2}{\sigma_2^2} + \frac{(\Delta - 1)^2}{\sigma_\Delta^2} , \quad (4)$$

where $\sigma_i = \rho_i m_i$; $i = 1, 2$ are the independent standard errors in the measurements. This equation clearly identifies each contribution to the overall uncertainty.

The desired distribution for x is obtained by integrating $p(x, \Delta | \mathbf{m})$ over the nuisance parameter Δ (a process called marginalization):

$$p(x | \mathbf{m}) = \int_{-\infty}^{\infty} p(x, \Delta | \mathbf{m}) \, d\Delta . \quad (5)$$

The analytic result for this integral is identical to the PPP answer, Eq. (1). We refer to this approach as Method A.

Probabilistic model

The above analysis effectively hides the source of the correlation between the uncertainties in the measurements. A detailed probabilistic formulation helps elucidate what is going on. We first treat the case in which each measurement is affected by a correlated additive offset, as could arise from a background subtraction.

We consider the joint probability of $x = x_1 = x_2$ and Δ, a systematic additive offset in x, which is uncertain (e.g., $x = m_1 + \Delta$). We use Bayes law [9] to obtain the posterior distribution, $p(x, \Delta | \mathbf{m}) = p(\mathbf{m} | x, \Delta) p(x, \Delta) = p(\mathbf{m} | x, \Delta) p(x) p(\Delta)$, where $p(\mathbf{m} | x, \Delta)$ is the likelihood, and $p(x)$ and $p(\Delta)$ are the priors on x and Δ. We assume $p(x)$ is uniform, indicating no prior information about x, and $p(\Delta)$ is normally distributed with its peak at $\Delta = 1$, and a standard error of $\sigma_\Delta = \sigma_c = 0.2$, to reflect what we know from the PPP statement.

Plausible experimental scenario

Suppose that to measure a cross section we must work in an experimental environment in which the background grows linearly in time, as shown in Fig. 2. This background could be caused by increasing activation of the apparatus, for example. Let us say that the measured rates m_1 and m_2 are made in the intervals shown, and that the background function is measured at some other time.

Knowing that the background always increases linearly from $t = 0$, we would estimate a background for the m_1 measurement to be 1.5 times that for m_2. Clearly this scenario could lead to a "fully correlated" contribution to the covariance with a ratio of 1.5^2:1.

In this situation, the background subtraction leads an additive correlated uncertainty, and the PPP answer of 0.882 is appropriate.
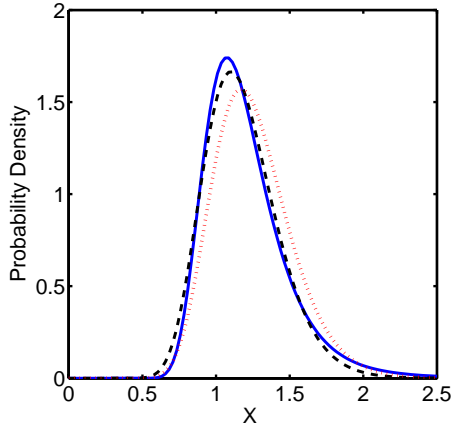
FIGURE 3. Probability density functions for methods B (solid line; normalization error), C (dashed; normalization error, log-normal prior), and D (dotted; logarithmic analysis).

ALTERNATIVE APPROACHES

In this section we offer several other interpretations of the stated facts. In the context of nuclear physics, a high degree of correlation might come from an overall normalization uncertainty, common to both measurements.

Normalization error

We use a probabilistic analysis to include a common error that arises from uncertainty in normalization. We assume the measurements must be divided by a normalization factor c to get $x = m/c$. Assuming normal distributions for the likelihood and for the uncertainty in c, the negative logarithm of the posterior distribution is

$$2\varphi = \frac{(cx - m_1)^2}{\sigma_1^2} + \frac{(cx - m_2)^2}{\sigma_2^2} + \frac{(c - 1)^2}{\sigma_c^2}, \quad (6)$$

written as function of cx and c. Transforming to variables x and c, we must divide the pdf by $|J|$, where J is the Jacobian for the transformation, i.e. the determinant of the first derivatives of the new variables with respect to the old ones; $J = 1/c$. Thus, $p(x,c) \propto |J|^{-1} \exp(-\varphi)$. We remove the nuisance parameter c by numerically integrating $p(x,c)$ over c. Figure 3 shows the result

The distribution is clearly not symmetric, and therefore not normal. For asymmetric distributions, a better estimate for x than $\hat{x}_{max}$ is the posterior mean $\hat{x}_{mean} = \langle x \rangle = \bar{x}$, where $\langle \rangle$ stands for averaging over the posterior distribution. The variance is estimated from the second moment around $\bar{x}$: $var(x) = \langle (x - \bar{x})^2 \rangle$.

The results for this approach, which we refer to as Method B, are $\hat{x}_{max} = 1.074$, $\hat{x}_{mean} = 1.200$, and $\sigma_x =$

$\sqrt{var(x)} = 0.276$. This analysis was offered as the best solution to the PPP by Don Smith [8] (p. 205ff). A similar approach is taken in [10]. Zhao and Perey [2] suggest a similar approach, but use the standard nonlinear least-squares technique to obtain $x = 1.15$ and $\sigma_x = 0.24$.

Other approaches

The normalization factor is a scale variable. If we had no knowledge of its approximate size, and wanted to use a noninformative prior to capture that lack of knowledge, it would be appropriate to use a uniform distribution in its logarithm. When transformed into a distribution in c, one gets $p(c) \propto c^{-1}$, which is called Jeffrey's prior. See [9] for a more complete argument involving the use of the maximum entropy principle. In this case, the scaling factor is taken to be unity with a stated uncertainty of $\sigma_c = 0.2$. To provide a smooth transition to the Jeffrey's prior in the limit of $\sigma_c \to \infty$, it seems appropriate to use a normal distribution in $\log(c)$, or log-normal in c.

Equation (6) now becomes

$$2\varphi = \frac{(cx - m_1)^2}{\sigma_1^2} + \frac{(cx - m_2)^2}{\sigma_2^2} + \frac{\log^2(c)}{\sigma_c^2}. \quad (7)$$

The Jacobian in this case is $J = 1$. Figure 3 shows the resulting posterior distribution, which is not very different from the previous result. The results for this approach, which we refer to as Method C, are $\hat{x}_{max} = 1.101$, $\hat{x}_{mean} = 1.177$ and $\sigma_x = 0.253$.

Another approach to coping with the PPP effect is to take the logarithms of the data [11]. Then multiplicative (i.e., normalization) uncertainties become additive. If the likelihoods are assumed to be normal distributions in $\log(x)$, they will be log-normal distributions in x with standard deviations of $\rho_1 = \rho_2 = 0.1$, and

$$2\varphi = \frac{(\log(x) - \log(m_1) - r)^2}{\rho_1^2}$$
$$+ \frac{(\log(x) - \log(m_2) - r)^2}{\rho_2^2} + \frac{r^2}{\sigma_r^2}. \quad (8)$$

The Jacobian is $J = x$. Figure 3 shows the resulting posterior distribution, which differs only slightly from the previous two distributions. The results for this approach, which we refer to as Method D, are $\hat{x}_{max} = 1.171$, $\hat{x}_{mean} = 1.252$ and $\sigma_x = 0.267$.

Chiba and Smith [12] suggest another approach in which the 20% correlated error is applied to the inferred value of x, not to each measurement. This assumption may be appropriate, depending the experimental situation. An iterative procedure yields the result $\hat{x} = 1.250$ and $\sigma = 0.28$.
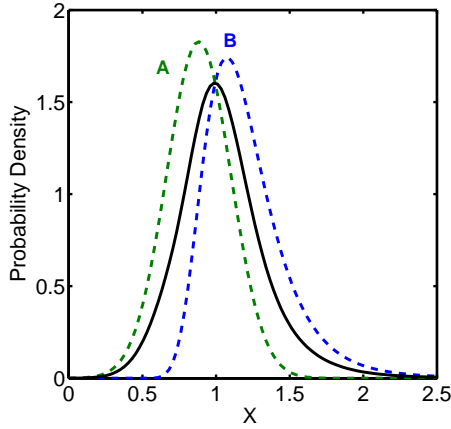
FIGURE 4. (solid line) Probability density function from combining two equally likely models, A and B (dashed lines).

TABLE 1. Results for various methods.

Method	$\hat{x}_{max}$	$\hat{x}_{mean}$	σ_x
A	0.882	0.882	0.228
B	1.074	1.200	0.276
C	1.101	1.177	0.253
D	1.171	1.252	0.267
A + B	0.991	1.041	0.295

Model uncertainty

The conclusion from the above discussion is that, without further information, we do not know which analysis model to use. This uncertainty can be handled in the probabilistic framework [9] as follows: $p(x|\mathbf{m}) = \sum_k p(x, M_k|\mathbf{m}) = \sum_k p(x|M_k, \mathbf{m})p(M)$, where $p(M_k)$ is the prior on model k. If we take models A and B to be equally likely, and ignore the other models, the result is $\frac{1}{2}p(x|M_A, \mathbf{m}) + \frac{1}{2}p(x|M_B, \mathbf{m})$, which is shown in Fig. 4. From the first and second moments of this distribution, the estimated value of x is 1.041 ± 0.295.

CONCLUSION

From Table 1 we see that the result given in the PPP (Method A) stands out because of its implicit assumption that the correlation contribution comes from an additive effect. We suggested an experimental situation in which this result is appropriate. The remaining methods all treat the correlation effect as multiplicative, in one way or another. Figure 3 shows that their posterior distributions do not significantly differ relative to their width.

The choice of which analysis approach to use should be made on the basis of one's best knowledge of how the measurements were made and the sources of their stated uncertainties. Full resolution of the PPP dilemma is only possible through improved knowledge of how the uncertainties contribute to the measurements. We argue that if one does not know which model to use, it is reasonable use all of them and average their posteriors, thus increasing the uncertainty in the answer.

We close with a plea to experimentalists to provide as many experimental details as possible when they report their results. Without such details, analysts will face the kind of ambiguity posed by Peelle's Pertinent Puzzle, necessarily inflating the uncertainty in the final result.

REFERENCES

1. Peelle, R. W., *Peelle's Pertinent Puzzle*, informal ORNL memorandum, October 13, 1987.
2. Zhao, A., and Perey, F. G., "Generation of covariance matrix for derived quantities and their least square fitting," in *Int. Symp. Nucl. Data Eval. Methodology*, edited by C. L. Dunford, World Scientific, New Jersey, 1993, pp. 657–662.
3. Chiba, S., and Smith, D. L., *J. Nucl. Sci. Tech.*, **31**, 770–781 (1994).
4. Oh, S.-Y., "Box-Cox transformation for resolving Peelle's pertinent puzzle in curve fitting," IAEA, Vienna, 2003, pp. 146–158, iNDC(NDS)-438.
5. Pronyaev, V. G., "Test and intercomparisons of data fitting with general least squares code GMA versus Bayesian code GLUCS," IAEA, Vienna, 2003, pp. 159–171, iNDC(NDS)-438.
6. Smith, D. L., and Pronyaev, V. G., "Update of GMA code to solve the PPP problem (technically)," IAEA, Vienna, 2004, pp. 333–342, INDC(NDS)-453.
7. Bevington, P. R., and Robinson, D. K., *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, Boston, 1992.
8. Smith, D. L., *Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology*, vol. 8 of *OECD Nucl. Energy Agency Nucl. Data Comm. Series*, Amer. Nucl. Soc., LaGrange Park, 1991.
9. Sivia, D. S., *Data Analysis - A Bayesian Tutorial*, Clarendon, Oxford, 1996.
10. d'Agostini, G., *Bayesian Reasoning in Data Analysis: A Critical Introduction*, World Scientific, New Jersey, 2003.
11. Kawano, T. et al., *J. Nucl. Sci. Tech.*, **37**, 327–334 (2000).
12. Chiba, S., and Smith, D. L., A suggested procedure for resolving an anomaly in least-squares analysis known as Peelle's pertinent puzzle and the implications for nuclear data evaluation, Tech. Rep. ANL/NDM-121, Argonne Nat'l. Lab. (1991).