LA-UR-21-28393
Accepted Manuscript

Analyzing count data with measurement error

Hamada, Michael Scott
Casleton, Emily Michele
Osthus, David Allen
Weaver, Brian Phillip
Steiner, Stefan H.

To be published in: Quality and Reliability Engineering International

DOI to publisher's version: 10.1002/qre.3078

Permalink to record:
http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-21-28393

Received: 9 September 2021 | Revised: 9 November 2021 | Accepted: 12 January 2022

DOI: 10.1002/qre.3078

RESEARCH ARTICLE

WILEY

Analyzing count data with measurement error

Michael S. Hamada[1] | **Emily M. Casleton**[1] | **Dave Osthus**[1] | **Brian P. Weaver**[1] | **Stefan H. Steiner**[2]

[1] Statistical Sciences Group, Los Alamos National Laboratory, Los Alamos, NM, USA

[2] Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada

Correspondence
Michael S. Hamada, Statistical Sciences Group, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
Email: hamada@lanl.gov

Abstract

In this article, we analyze observed count data such as the number of defects in a steel product where the observed counts are the true counts measured with errors. We account for the measurement error by using a measurement error model based on a latent lognormal (LLN) distribution. We consider making inference about a single population (e.g., from samples of a production lot) and a regression model (e.g., from runs of a designed experiment), where the measurement system properties are known, that is, the parameters of the LLN distribution are known. Then, we consider simultaneous inference for the single population and regression model as well as the measurement system. We demonstrate the proposed methodology with both simulated and real observed counts.

KEYWORDS

Bayesian inference, designed experiment, gauge R & R study, latent lognormal distribution, Markov chain Monte Carlo, Poisson, Poisson regression, regression, single population

1 | INTRODUCTION

Count measurements arise in many applications, for example, the number of paint imperfections on an automobile hood or steel product defects from inclusion of nonmetallic materials during production. Later we report on an application that counted the number of bubbles blown according to various bubble recipes.

In this article, we consider the analysis of observed count data where the true counts are a sample of the population. For example, the true counts could be Poisson-distributed but the observed counts are measured with error. Here, we want to make inference about the population. We also consider the relationship between a covariate x and the true counts through a Poisson regression model or more generally in a designed experiment when there are multiple covariates, that is, a vector $\mathbf{x}$ of covariates corresponding to main effects and interactions. Again we consider the analysis of observed count data, that is, the true counts measured with error. In this article, we use the recent count measurement error model proposed in Osthus et al.[1] This article is the first to show how inference about a single population or regression model can be done with a known count measurement system based on the this count measurement error model. This article is also the first to show how inference about a single population or regression model can be done simultaneously while also assessing the count measurement system.

An outline of the paper is as follows. In Section 2, we consider Poisson models for the true counts. Then, we consider a model for the observed counts where the true counts are measured with error in Section 3. The measurement error model allows for under- and over-dispersion by a deflation/inflation factor where deflation and inflation correspond to under-dispersion and over-dispersion, respectively. Then, in Section 4, we briefly discuss Bayesian inference for observed counts. Next, we consider a number of cases. First, we consider a single population with known measurement system properties in Section 5. Then, we consider regression with known measurement system properties in Section 6. Next, in Section 7, we consider a gauge R & R study in which we simultaneously consider a single population and assess a measurement system (when the measurement system properties are unknown). Then, in Section 8, we consider a mixture experiment in which we simultaneously consider a regression model and assess a measurement system. Finally, we conclude with a discussion in Section 9.

2 | MODELS FOR TRUE COUNTS

For a sample of true counts θ_i, we assume that the population of true counts is described by a Poisson distribution with mean θ.

In a Poisson regression model, the true count θ_i is related to the covariate x_i by $\theta_i \sim Poisson(\lambda_i)$, where mean $\lambda_i = \exp(\beta_0 + \beta_1 x_i)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1)$ is a vector of regression coefficients. There can be multiple covariates such as in a designed experiment corresponding to main effects and interaction effects. Here θ_i is the true count for the ith run, $\mathbf{x}_i$ is a vector of covariates for the ith run, and $\boldsymbol{\beta}$ is the vector of regression coefficients so that $\lambda_i = \exp(\mathbf{x}_i\boldsymbol{\beta})$. If the experiment is replicated, then θ_{ij} is the true count of the jth replicate at the ith run, and $\theta_{ij} \sim Poisson(\lambda_i)$.

Let y_i be the observed count, that is, the true count measured with error, corresponding to the true count θ_i. In a designed experiment with replicates, let y_{ij} be the observed count of the jth replicate at the ith run corresponding to the true count θ_{ij}. Later, when we consider experiments that allow the simultaneous inference of a single Poisson population or a Poisson regression model as well as the count measurement system, the structure of the observed counts y will be more complicated and therefore have more indices.

3 | A MODEL FOR OBSERVED COUNTS

We consider a measurement process where the measurements are counts (non-negative integers). More concretely, let $y_i \in \{0, 1, 2, ...\}$ be the observed count where θ_i denotes the true count. Osthus et al.[1] consider measurement system assessment (MSA) for counts and are motivated by the Poisson distribution for observed counts Y, where $Var(Y) = E(Y)$. To be more flexible, Osthus et al.[1] consider

$$StdDev(Y) \approx \alpha E(Y), \tag{1}$$

where $\alpha < 1$ deflates the standard deviation and $\alpha > 1$ inflates the standard deviation so that α is a deflation/inflation factor. Note that Osthus et al.[1] is the first article to consider count data MSA. Following Osthus et al.[1], we propose a latent lognormal (LLN) model, $Y \sim LatentLognormal(\mu, \sigma^2)$, where Y is the discretized $\tilde{Y}$ and $\tilde{Y} \sim Lognormal(\mu, \sigma^2)$, that is, $\tilde{Y}$ has a lognormal distribution. We discretize $\tilde{Y}$ by specifying the interval cut points as $\{0, 0.5, 1.5, 2.5, 3.5, ...\}$, that is, $Y = 0$ if $\tilde{Y} < 0.5$, $Y = 1$ if $0.5 \leq \tilde{Y} < 1.5$, and so on. In Equation (1), the relationship is approximate because $StdDev(\tilde{Y}) = \alpha_{\tilde{Y}} E(\tilde{Y})$, that is, on the $\tilde{Y}$ scale. This relationship is nearly exact because the $E(Y)$ and $StdDev(Y)$ are nearly equal to $E(\tilde{Y})$ and $StdDev(\tilde{Y})$, respectively. For example, for $Y \sim LatentLognormal(50, 1^2)$, $E(Y) = 82.2497$, $E(\tilde{Y}) = 82.2493$, $StdDev(Y) = 107.4670$, and $StdDev(\tilde{Y}) = 107.4665$.

Now, let us use the LLN model to specify the y_i distribution as

$$y_i | \mu_i, \sigma_i \sim LatentLognormal(\mu_i, \sigma_i^2), \tag{2}$$

for

$$\mu_i = \ln\left(\sqrt{\frac{(\theta_i \gamma)^2}{1 + (\frac{\alpha_{\tilde{Y}}}{\gamma})^2}}\right) \quad \text{and}$$

$$\sigma_i^2 = \ln\left(1 + (\frac{\alpha_{\tilde{Y}}}{\gamma})^2\right), \tag{3}$$

where θ_i is the true count and γ is the multiplicative bias for the operator making the measurement. From Equations (2) and (3), we have

$$E(\tilde{y}_i|\mu_i,\sigma_i) = \theta_i\gamma \quad \text{and}$$

$$StdDev(\tilde{y}_i|\mu_i,\sigma_i) = \alpha_{\tilde{Y}}\theta_i. \tag{4}$$

We see that $\alpha_{\tilde{Y}}$ is a deflation/inflation factor.

If the MSA has been done and the particular operator that did the measurement y_i is known, say the kth operator, then γ_k would replace γ in Equation (3). Otherwise,

$$\gamma_k|\sigma_\gamma \sim Lognormal(0,\sigma_\gamma^2), \tag{5}$$

where the operators used in the gauge R & R study to perform the MSA are assumed to be a random sample from a population of operators whose γ_k follow a lognormal distribution. Note that the median of γ_k is 1 so that γ_k under- or over-estimates θ_i if $\gamma_k < 1$ or $\gamma_k > 1$, respectively.

In the case where there are replicates θ_{ij}, all indexing with i in Equations (2)–(4) is replaced with the indexing ij. If the kth operator took the measurement, then the notation y_{ijk} is used; if no replicates were taken, we use the notation $y_{i\cdot k}$, where $\cdot$ indicates that the second index j is not used. If the operator is not known, then we use $y_{ij\cdot}$. If there are repeat measurements on the ith true count, then the index l is used, so that y_{ijkl} denotes the corresponding observed count. Again we use the $\cdot$ notation where the index is not applicable, for example, with no replicates, but where the operator is known and there are repeats, $y_{i\cdot kl}$ denotes the observed counts.

4 | BAYESIAN INFERENCE

In the subsequent examples, we use Bayesian inference through Bayes' Theorem (Gelman et al.[2]) to obtain posterior distributions for the unknown model parameters. For example, if the measurement system parameters are known (which we refer to as a known measurement system), we treat them as constants and they are not updated in the analysis of the example data. Modern Bayesian computing through Markov chain Monte Carlo or MCMC has made it possible to sample from the posterior distribution (Gelman et al.[2]). We implement the various data analyses presented in the examples using JAGS (Plummer[3]) called by the R (R Core Team[4]) package rjags to obtain draws or samples from the parameter posterior distributions. In the examples, we use at least 1000 burnin draws and at least 100,000 subsequent draws with a thinning of every 10 draws. Diagnostics (Gelman and Rubin's convergence diagnostic[5]) and plots of these draws suggest convergence, that is, that these draws are from the appropriate posterior distributions. In the examples, we provide the parameter prior distributions that we use.

5 | SINGLE POISSON POPULATION WITH KNOWN MEASUREMENT SYSTEM PROPERTIES

In this section, we consider a single population with known measurement system. We simulate a sample of $N = 50$ true counts θ_i distributed as $Poisson(\theta)$ with $\theta = 50$. We assume that the measurement system properties are known: $\gamma_k = \gamma = 1$ for all the operators, that is, the measurement system is unbiased, and $\alpha_{\tilde{Y}} = 1$. We use the known measurement system to simulate observed counts y_i from the true counts. See Figure 1 that plots the true counts θ_i and the observed counts y_i. For the prior distribution for θ, we use $\theta \sim Lognormal(3.91, 1^2)$; the prior median and 0.95 probability interval are 49.90 and (7.029, 354.236). We see that this prior is quite diffuse. In practice, one can use an appropriate diffuse prior or one that is more informative if knowledge is available to choose a narrower prior.

See Table 1 that displays posterior summaries (50, 2.5, 97.5 percentiles) when the true counts are analyzed, when the observed counts are analyzed but the measurement error is ignored, and when the observed counts are analyzed accounting for the measurement error. When the measurement error is ignored, the results in the second row are very similar to those for the true counts in the first row. However, the 0.95 probability interval (2.5 and 97.5 percentiles) is too narrow as

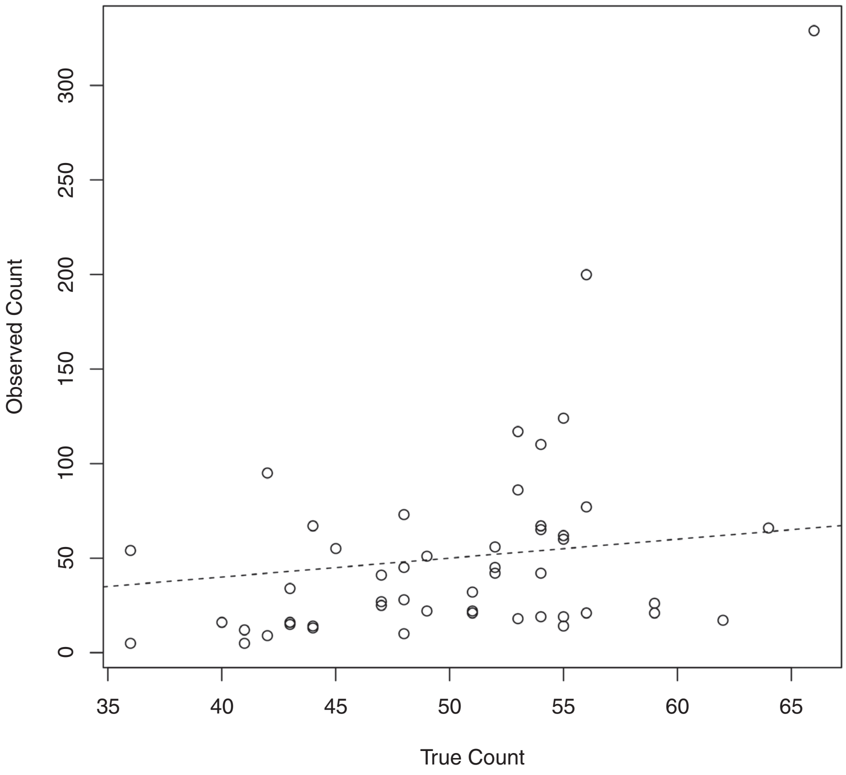

FIGURE 1 Single Poisson population with $\theta = 50$, true count versus observed count from measurement system with $\gamma = 1$, $\alpha_{\tilde{Y}} = 1$ for sample size $N = 50$

TABLE 1 Posterior summaries (50, 2.5, 97.5 percentiles) for θ, single Poisson population with mean $\theta = 50$

Case	50%	2.5%	97.5%
True counts	50.120	48.182	52.118
$\gamma = 1, \alpha_{\tilde{Y}} = 1$			
Ignore meas. err.	50.203	48.261	52.194
Account meas. err.	48.213	38.347	60.745
$\gamma = 0.75, \alpha_{\tilde{Y}} = 1$			
Ignore meas. err.	37.660	35.958	39.394
Account meas. err.	48.491	36.648	64.126

seen by the results when the measurement error is accounted for in the third row. In this case, we see that the estimate of θ, say the 50 percentile, is not affected by ignoring the measurement error.

Next, we assume that the measurement system is biased low: $\gamma_k = 0.75$ for all the operators and $\alpha_{\tilde{Y}} = 1$. When the measurement error is ignored, we see in the fourth row of Table 1 that inference for θ is biased low and also that the 0.95 probability interval is too narrow as seen by the results when the measurement error is accounted for in the fifth row. We also see that inference for θ is not biased when the measurement error is accounted for in the fifth row.

Instead of the one parameter Poisson population, consider a LLN population with $\theta = 50$ and $\alpha_\theta = 0.25$ (γ_θ is set to 1 and is known) and LLN measurement error with $\gamma_{\tilde{Y}} = 1$ and $\alpha_{\tilde{Y}} = 1$. As before, we simulate $N = 50$ true counts and simulate observed counts with measurement error.

See Table 2 that displays posterior summaries (50, 2.5, 97.5 percentiles) when the true counts are analyzed, when the observed counts y_i are analyzed but the measurement error is ignored, and when the observed counts are analyzed accounting for the measurement error. When the measurement error is ignored, the estimate for θ, the 50 percentile in the third row for θ is similar to that for the true counts in the first row; the width of the 0.95 probability interval is slightly wider than that when the measurement error is accounted for. However, the inference for α_θ is very biased with the median much greater than the true value as seen in the results in the fourth row as compared with that for the true counts in the second row. When the measurement error is accounted for in the sixth row, inference for α_θ is similar to that

TABLE 2 Posterior summaries (50, 2.5, 97.5 percentiles) for θ, single latent lognormal population with $\theta = 50$, $\alpha_\theta = 0.25$, $\gamma_{\tilde{Y}} = 1$, and $\alpha_{\tilde{Y}} = 1$

Case	Parameter	50%	2.5%	97.5%
True counts	θ	49.513	46.206	53.334
	α_θ	0.253	0.208	0.317
Ignore meas. err.	θ	51.053	39.786	69.724
	α_θ	1.015	0.789	1.405
Account meas. err.	θ	52.536	40.933	69.914
	α_θ	0.210	0.009	0.661

for the true counts albeit more uncertain, that is, the 0.95 probability interval is wider. The inference for θ is also more uncertain than that for the true counts, another impact of the measurement error.

6 | POISSON REGRESSION WITH KNOWN MEASUREMENT SYSTEM PROPERTIES

In this section, we consider a regression with counts, that is, Poisson regression, with known measurement system. Wu and Hamada[6] present an experiment on a wave soldering process that reports the solder-joint defects on circuit boards made by the process. Because we do not know the true values of the effects nor anything about the measurement system, we decided to simulate true and observed counts with measurement error to assess the impact of measurement error in the regression setting much like we did for the single population case. We use a 2^4 design for four factors replicated five times so that there are $N = 80$ observations and consider all 16 effects: INT for intercept, (1, 2, 3, 4) for the four main effects, (12, 13, 14, 23, 24, 34) for the six two-factor interactions, (123, 124, 134, 234) for the four three-factor interactions, and 1234 for the four-factor interaction. For the ith run mean, we have $\lambda_i = \exp(\mathbf{x}_i\boldsymbol{\beta})$ and $\mathbf{x}_i$ are the associated covariates for the ith run, where $\boldsymbol{\beta} = (2, 1, .75, .5, 0, .5, .25, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ so that INT, 1, 2, 3, 12, 13 are active. The prior distribution for $\boldsymbol{\beta}$ is $\beta_0 \sim Normal(3, \sigma^2_{\beta0} = 1^2)$ for INT and $\beta_i \sim Normal(0, \sigma^2_{\beta0} = 1^2)$ for $i = 2, \ldots, 16$, that is, associated with effects (1, 2, ..., 1234). We assume that the measurement system is known: $\gamma_k = \gamma = 1$ for all the operators, that is, the measurement system is unbiased, and $\alpha_{\tilde{Y}} = 1$.

Table 3 displays the posterior 50, 2.5, and 97.5 percentiles of the effects for the three cases: true counts, observed counts but ignore measurement error (i.e., treat as true counts), and observed counts where measurement error is accounted for.

TABLE 3 Posterior summaries (50, 2.5, 97.5 percentiles) for Poisson regression example with unbiased measurement error

		True counts			Ignore meas. err.			Account meas. err.		
Effect	True Value	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
INT	2	1.910	1.786	2.026	1.577	1.426	1.714	1.999	1.788	2.201
1	1	1.059	0.942	1.183	1.104	0.965	1.254	1.021	0.823	1.227
2	0.75	0.849	0.731	0.974	0.799	0.659	0.947	0.756	0.554	0.961
3	0.5	0.456	0.335	0.580	0.485	0.344	0.632	0.455	0.251	0.658
4	0	0.006	−0.115	0.128	−0.025	−0.169	0.118	−0.022	−0.227	0.184
12	0.5	0.418	0.294	0.536	0.463	0.316	0.604	0.472	0.268	0.679
13	0.25	0.269	0.147	0.390	0.258	0.114	0.403	0.272	0.068	0.480
14	0	0.007	−0.114	0.129	−0.002	−0.146	0.142	−0.019	−0.222	0.187
23	0	0.067	−0.057	0.188	0.124	−0.020	0.266	0.108	−0.097	0.313
24	0	0.076	−0.044	0.200	0.067	−0.076	0.211	0.034	−0.172	0.240
34	0	0.012	−0.108	0.133	0.049	−0.094	0.194	0.008	−0.199	0.211
123	0	−0.052	−0.173	0.072	−0.099	−0.244	0.044	−0.099	−0.306	0.102
124	0	−0.095	−0.220	0.026	−0.119	−0.266	0.022	−0.087	−0.292	0.119
134	0	−0.079	−0.201	0.041	−0.101	−0.246	0.041	−0.072	−0.274	0.132
234	0	−0.047	−0.169	0.074	0.003	−0.143	0.145	0.019	−0.187	0.223
1234	0	0.094	−0.025	0.218	0.064	−0.078	0.210	0.043	−0.160	0.251

TABLE 4 Posterior summaries (50, 2.5, 97.5 percentiles) for Poisson regression example with biased measurement error

		True counts			Ignore meas. err.			Account meas. err.		
Effect	**True Value**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**
INT	2	1.910	1.786	2.026	1.102	0.904	1.279	2.04	1.808	2.266
1	1	1.059	0.942	1.183	1.148	0.968	1.346	1.021	0.792	1.255
2	0.75	0.849	0.731	0.974	0.815	0.632	1.007	0.722	0.493	0.953
3	0.5	0.456	0.335	0.580	0.493	0.307	0.687	0.441	0.208	0.673
4	0	0.006	−0.115	0.128	−0.018	−0.210	0.169	−0.009	−0.233	0.221
12	0.5	0.418	0.294	0.536	0.422	0.230	0.607	0.455	0.224	0.685
13	0.25	0.269	0.147	0.390	0.230	0.038	0.416	0.253	0.018	0.482
14	0	0.007	−0.114	0.129	0.007	−0.180	0.198	−0.014	−0.243	0.218
23	0	0.067	−0.057	0.188	0.154	−0.036	0.343	0.133	−0.097	0.360
24	0	0.076	−0.044	0.200	0.059	−0.126	0.250	0.017	−0.208	0.250
34	0	0.012	−0.108	0.133	0.067	−0.117	0.260	0.002	−0.230	0.231
123	0	−0.052	−0.173	0.072	−0.105	−0.295	0.085	−0.088	−0.322	0.141
124	0	−0.095	−0.220	0.026	−0.129	−0.320	0.056	−0.089	−0.321	0.142
134	0	−0.079	−0.201	0.041	−0.134	−0.328	0.048	−0.080	−0.309	0.154
234	0	−0.047	−0.169	0.074	−0.037	−0.229	0.145	0.003	−0.229	0.233
1234	0	0.094	−0.025	0.218	0.120	−0.063	0.314	0.071	−0.165	0.303

When the measurement error is ignored, the results are similar to those for true counts, except that the posterior for INT is somewhat lower although this has no impact on determining active effects. However, the uncertainty is too small (i.e., the 0.95 probability intervals are too narrow) as compared with the results when the measurement error is accounted for; also the results for INT agree with those for true counts when the measurement error is accounted for. The main impact of ignoring the measurement error is that more effects could be identified as active because the uncertainty is too small so that 0.95 probability intervals might not contain zero, whereas those when the measurement error is accounted for could contain zero.

Next we consider the case when the measurement system is biased: $\gamma_k = \gamma = 0.75$ for all the operators and $\alpha_{\tilde{Y}} = 1$. Table 4 displays the posterior 50, 2.5, and 97.5 percentiles of the effects for the three cases: true counts (i.e., same as in Table 3), observed counts but ignore measurement error (i.e., treat as true counts), and observed counts where measurement error is accounted for. The Table 4 results are similar to those of Table 3 except that the posterior for INT is much lower although this has no impact on determining active effects.

Like the single population case, we consider a LLN distribution for the true counts with mean λ_i and $\alpha_\theta = 0.25$. The measurement error is also a LLN distribution with $\gamma_k = \gamma = 1$ and $\alpha_{\tilde{Y}} = 1$. Table 5 displays the posterior 50, 2.5, and 97.5 percentiles of the effects for the three cases: true counts, observed counts but ignore measurement error (i.e., treat as true counts), and observed counts where measurement error is accounted for. When the measurement error is ignored, we see that α_θ is overestimated. Unlike the previous results, the 0.95 probability intervals are similar in width to those when the measurement error is accounted for; they are wider when the measurement error is accounted for than for true counts as expected.

7 SIMULTANEOUS INFERENCE FOR A SINGLE POISSON POPULATION AND MSA

In this section, we consider simultaneous inference for a single population with unknown measurement system. A gauge R & R study is an experiment that provides data to perform MSA. Typically, there is a random sample of n_P parts, a random sample of n_K operators and each operator measures each part n_r times, that is, n_r repeats.[7] Let θ_i be the true counts and $y_{i \cdot kl}$ are the observed counts for $i = 1, \ldots, n_P$, $k = 1, \ldots, n_K$, and $l = 1, \ldots, n_r$. As in the Section 5, we assume that $\theta_i \sim Poisson(\theta)$. The $\gamma_k \sim Lognormal(0, \sigma_\gamma^2)$ are the operator biases and $\alpha_{\tilde{Y}}$ is the deflation/inflation factor that the observed count mean and variance depend on $\mu_{i \cdot k}$ and $\sigma_{i \cdot k}$ in Equation (3) replacing γ with γ_k as given in Equation (4),

TABLE 5 Posterior summaries (50, 2.5, 97.5 percentiles) for latent lognormal regression example with unbiased measurement error

		True counts			**Ignore meas. err.**			**Account meas. err.**		
Effect	**True Value**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**
$\alpha_{\tilde{Y}}$	0.25	0.264	0.216	0.329	1.156	0.825	1.852	0.365	0.018	1.037
INT	2	2.010	1.949	2.073	2.052	1.762	2.474	2.056	1.787	2.445
1	1	1.003	0.943	1.065	0.970	0.701	1.233	0.965	0.702	1.231
2	0.75	0.718	0.658	0.779	0.717	0.450	0.981	0.718	0.454	0.986
3	0.5	0.452	0.391	0.513	0.316	0.050	0.579	0.317	0.053	0.582
4	0	0.028	−0.032	0.090	−0.132	−0.398	0.130	−0.136	−0.398	0.129
12	0.5	0.480	0.419	0.541	0.420	0.154	0.681	0.419	0.155	0.680
13	0.25	0.223	0.162	0.284	0.224	−0.046	0.490	0.223	−0.047	0.485
14	0	0.017	−0.043	0.078	0.107	−0.160	0.377	0.104	−0.161	0.366
23	0	−0.005	−0.066	0.056	0.143	−0.125	0.405	0.145	−0.119	0.403
24	0	−0.003	−0.064	0.058	0.063	−0.205	0.328	0.065	−0.199	0.338
34	0	−0.006	−0.067	0.055	0.225	−0.044	0.488	0.221	−0.042	0.483
123	0	0.000	−0.062	0.061	0.008	−0.257	0.275	0.009	−0.255	0.277
124	0	0.030	−0.031	0.091	0.061	−0.205	0.326	0.058	−0.217	0.322
134	0	−0.024	−0.085	0.038	−0.123	−0.393	0.140	−0.124	−0.394	0.145
234	0	0.033	−0.027	0.095	0.124	−0.144	0.389	0.123	−0.139	0.393
1234	0	0.006	−0.055	0.067	−0.110	−0.375	0.155	−0.112	−0.378	0.152

where $\gamma = \gamma_k$ is associated with observed counts $y_{i\cdot kl}$. That is, θ describes the single population and γ_k, σ_γ, and $\alpha_{\tilde{Y}}$ describe the measurement system.

We could also add a non-negative multiplicative part by operator effects δ_{ik} to $\mu_{i\cdot k}$ and $\sigma^2_{i\cdot k}$ in Equation (3) so that

$$\begin{aligned} E(\tilde{y}_{i\cdot kl}|\mu_{i\cdot k},\sigma_{i\cdot k}) &= \theta_i\gamma_k\delta_{ik} \text{ and} \\ StdDev(\tilde{y}_{i\cdot kl}|\mu_{i\cdot k},\sigma_{i\cdot k}) &= \alpha_{\tilde{Y}}\theta_i, \end{aligned} \tag{6}$$

where

$$\delta_{ik}|\sigma_\delta \sim Lognormal(0,\sigma^2_\delta). \tag{7}$$

Note that the median of δ_{ik} is 1 so that δ_{ik} also under- or over-estimates θ_i.

The prior distributions can be specified as:

$$\begin{aligned} &\theta \sim Lognormal(\mu_\theta,\sigma^2_\theta), \\ &\sigma_\gamma \sim HalfNormal(0,\sigma^2_{\gamma 0}), \\ &\sigma_\delta \sim HalfNormal(0,\sigma^2_{\delta 0}), \\ &\alpha_y \sim HalfNormal(0,\sigma^2_{\alpha_{\tilde{Y}}0}). \end{aligned} \tag{8}$$

We use the same sample of the Poisson population in Section 5, that is, $n_P = 50$, that was generated from $Poisson(50)$ so that $\theta_i \sim Poisson(\theta)$, where $\theta = 50$. For the measurement error, we use $\alpha_{\tilde{Y}} = 0.25$. Also, $\delta_{ij} = 1$ for all i and j, so that $\sigma_\delta = 0$. We also have $N_K = 5$ operators and $n_r = 3$ repeats. That is, each of five operators measure each of the $n_P = 50$ parts, $n_r = 3$ times. We set $\sigma_\gamma = 0.1$ and generated true $\boldsymbol{\gamma} = (0.978, 1.067, 1.264, 0.831, 1.047)$. For the prior distributions for θ, $\alpha_{\tilde{Y}}$, σ_γ, and σ_δ, we use $\mu_{\theta 0} = 3.91$, $\sigma_{\theta 0} = 1$, $\sigma_{\alpha_{\tilde{Y}}0} = 1$, $\sigma_{\gamma 0} = 1$, and $\sigma_{\delta 0} = 1$. Table 6 displays posterior distribution summaries (50, 2.5, and 97.5 percentiles). These results show that for the most part, the true parameter values are contained within the 0.95 probability intervals (2.5 and 97.5 percentiles). For example, the 0.95 probability interval contains the true value of θ, the mean of the Poisson population. The results for σ_δ also show that it is small. Note that Equation (5)

TABLE 6 Posterior summaries for Gauge R & R Study data

Parameter	True value	50%	2.5%	97.5%
θ	50.000	50.535	48.405	52.729
$\alpha_{\tilde{Y}}$	0.250	0.238	0.224	0.253
γ_1	0.978	0.959	0.927	0.991
γ_2	1.067	1.050	1.017	1.083
γ_3	1.264	1.234	1.198	1.270
γ_4	0.831	0.781	0.751	0.810
γ_5	1.047	1.032	0.994	1.069
σ_γ	0.100	0.212	0.107	0.634
σ_δ	0.000	0.034	0.005	0.058

TABLE 7 Bubble count data, five replicates, three operators

W	S	G	Soap Brand	Water Type	1	2	3
0.7	0.25	0.05	Joy	Spring	11	12	14
					13	14	14
					15	15	15
					9	13	13
					16	16	17
…	…	…	…	…	…	…	…
					…	…	…
					…	…	…
					…	…	…
					…	…	…
0.8	0.2	0	Joy	Tap	14	16	16
					12	11	13
					16	14	16
					12	13	16
					14	12	16

assumes that the γ_k are distributed about 0, that is, with no bias, but in a real data set, we would not be able to assess this assumption unless the true counts $\tilde{Y}$ were known. Nevertheless, the relative biases can still be assessed with the proposed methodology.

8 | SIMULTANEOUS INFERENCE FOR A DESIGNED EXPERIMENT AND MSA

Steiner et al.[8] present a mixture experiment whose response was the number of bubbles counted that were blown from different bubble recipes. In their analysis, they used the data from one operator; there were actually three operators. In the experiment, 48 bubble recipes were made and bubbles were blown five times from each recipe or five sets of bubbles per recipe. In other words, there were five replicates for each of the 48 runs where each operator measures it once, that is, a single repeat. The measured counts of the three operators from two of these runs are displayed in Table 7. For example, in the first replicate of the first run, the three operators counted 11, 12, and 14 bubbles, respectively. A bubble recipe is a mixture of soap (S), water (W), and glycerin (G) whose proportions sum to 1. There are two other factors, Soap Brand (Joy, Ivory) and Water Type (Spring, Tap); these factors are referred to as noise factors (and denoted by N_1 and N_2) because it is desirable to find a recipe robust across the level combinations of the noise factors.

For the jth replicate at the ith run, we assume the true count $\theta_{ij} \sim Poisson(\lambda_i)$. See Figure 2 that displays the run sample variance/run average versus run average plot for the replicates and shows that the average of the run sample variance/run

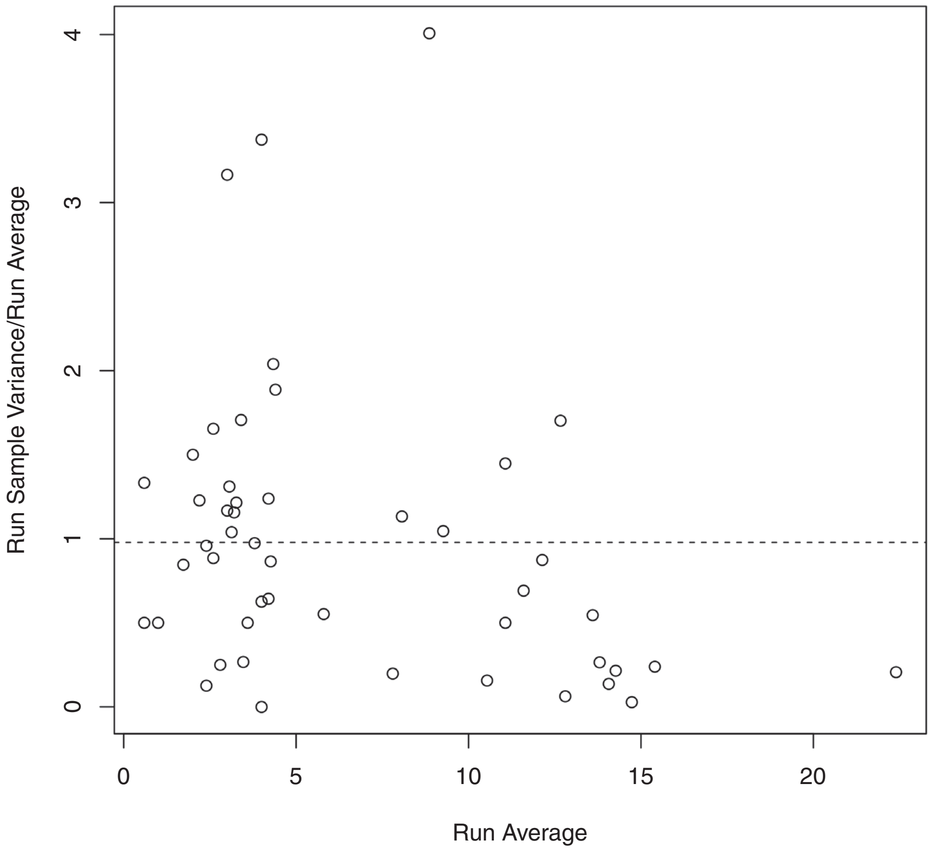

FIGURE 2 Bubble recipe experiment, run sample variance/run average versus run average

average 0.978 as a dashed line that is consistent with a Poisson (whose ratio would be equal to 1). For the run mean λ, we have $\lambda_i = \exp(\mathbf{x}_i\boldsymbol{\beta})$ and $\mathbf{x}_i$ are the associated covariates for the ith run; the names of the effects associated with these covariates are listed in Table 8. Note that there is no intercept (to avoid redundancy) because the experiment design is for a mixture experiment. For measurement error, the LLN model is assumed for the observed count, that is, $E(\tilde{Y}_{ijk}) = \gamma_k\theta_{ij}$ and $StdDev(\tilde{Y}) = \alpha_{\tilde{Y}}\theta_{ij}$ for the kth operator.

The prior distribution for $\boldsymbol{\beta}$ is $\beta_i \sim Normal(0, \sigma^2_{\beta 0} = 1^2)$. The prior distribution parameters for $\alpha_{\tilde{Y}}$ and σ_γ are $\sigma_{\alpha_{\tilde{Y}}0} = 1$ and $\sigma_{\gamma 0} = 1$. Note that no part by operator interactions δ_{ik} was considered in the LLN model because there were no repeats in this experiment; that is, there was a single repeat for each of the three operators. Table 8 displays posterior distribution summaries (50, 2.5, and 97.5 percentiles) for the MSA ($\alpha_{\tilde{Y}}, \sigma_\gamma, \gamma_1$-$\gamma_3$). These results show that the measurement system is quite good with both $\alpha_{\tilde{Y}}$ and σ_γ small. For the bubble recipe experiment, the 0.95 probability intervals show that W, S, G, WS, and WN_1 effects are active, that is, their 0.95 probability intervals do not contain 0.

9 | DISCUSSION

In this article, we proposed methodology for analyzing count data with measurement error using a measurement error model based on a LLN distribution. The measurement error model is flexible in that it can accommodate either under- or over-dispersion by a deflation/inflation factor. Through examples, we showed how inference for a single population or a regression model that arises in analyzing a factorial experiment can be done. Our analysis approach allows for the measurement system to have known properties or we can estimate the measurement system properties, provided that appropriate data are collected. Then simultaneous inference for the single population or regression model and the measurement system parameters can be performed. The examples showed that properly accounting for the measurement error provides better inference than otherwise can be obtained when the measurement error is ignored.

TABLE 8 Posterior distribution summaries (50, 2.5, and 97.5 percentiles) for the bubble recipe experiment

Effect	50%	2.5%	97.5%
$\alpha_{\tilde{Y}}$	0.200	0.187	0.215
γ_1	0.996	0.975	1.009
γ_2	1.000	0.984	1.017
γ_3	1.005	0.990	1.026
σ_γ	0.013	0.001	0.234
W	1.130	0.949	1.306
S	1.997	0.963	3.023
G	1.572	0.279	2.871
WS	2.499	0.856	4.155
WG	0.431	−1.230	2.088
SG	1.050	−0.865	2.977
WSG	0.697	−1.267	2.650
WN_1	−0.225	−0.406	−0.044
SN_1	−0.811	−1.840	0.222
GN_1	−0.601	−1.887	0.675
WSN_1	−1.502	−3.166	0.168
WGN_1	0.116	−1.532	1.756
SGN_1	−1.049	−2.921	0.864
$WSGN_1$	−0.695	−2.600	1.261
WN_2	0.044	−0.132	0.216
SN_2	−0.439	−1.454	0.580
GN_2	0.284	−1.001	1.582
WSN_2	0.014	−1.632	1.633
WGN_2	0.419	−1.266	2.084
SGN_2	−0.208	−2.123	1.678
$WSGN_2$	−0.147	−2.082	1.796
WN_1N_2	0.028	−0.148	0.209
SN_1N_2	0.015	−1.022	1.052
GN_1N_2	0.266	−1.024	1.527
WSN_1N_2	−0.592	−2.277	1.036
WGN_1N_2	−0.004	−1.685	1.610
SGN_1N_2	0.353	−1.543	2.263
$WSGN_1N_2$	0.207	−1.711	2.145

ACKNOWLEDGMENTS

We thank Cee Cee Essix for her support and encouragement. We also thank two anonymous referees whose insightful comments on an earlier version helped to improve the exposition of this article.

DATA AVAILABILITY STATEMENT

Data used in this article were simulated as described in this article or were provided in a table.

ORCID

Michael S. Hamada https://orcid.org/0000-0003-3206-1695

REFERENCES

1. Osthus D, Weaver BP, Casleton EM, Hamada MS, Steiner SH. On gauge R & R studies for counts. *Qual Eng*. 2021. https://doi.org/10.1080/08982112.2021.1957109.

2. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*. 3rd ed. Boca Raton: Chapman & Hall/CRC; 2013.
3. Plummer M. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Vienna, March 20–22; 2003.
4. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org. Accessed August 20, 2021.
5. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7:457–472.
6. Wu CFJ, Hamada MS. *Experiments: Planning, Analysis and Optimization*. 3rd ed. New York: John Wiley and Sons, Inc.; 2021.
7. Montgomery DC, Runger GC. Gauge capability and designed experiments: part I: basic methods. *Qual Eng*. 1993;6:115–135.
8. Steiner S, Hamada M, Giddings-White B, Kutsyy V, Mosesova S, Salloum G. A bubble mixture experiment project for use in an advanced design of experiments class. *J Stat Educ*. 2007;15(1).www.amstat.org/publications/jse/v15n1/steiner.html

AUTHOR BIOGRAPHIES

Michael S. Hamada is a Scientist and holds a Ph.D. in Statistics from the University of Wisconsin – Madison. He is a Fellow of the American Statistical Association, American Society for Quality, and Los Alamos National Laboratory. His research interests include design and analysis of experiments, measurement system assessment, quality control, and reliability.

Emily M. Casleton is a Scientist at Los Alamos National Laboratory and holds a Ph.D. in Statistics from Iowa State University. Her research interests include multisource data fusion, nonparametric Bayesian methods, and Bayesian design of experiments.

Dave Osthus is a Scientist at Los Alamos National Laboratory and holds a Ph.D. in Statistics from Iowa State University. His research interests include infectious disease forecasting, Bayesian methods for dependent data, and data visualization.

Brian P. Weaver is a Scientist and holds a Ph.D. in Statistics from Iowa State University. His research interests include measurement error, measurement system assessment, reliability, and astrostatistics.

Stefan H. Steiner is a Professor/Department Chair in the Department of Statistics and Actuarial Science. He holds a Ph.D. in Business Administration (Management Science/Systems) from McMaster University. His primary research interests include quality improvement, process monitoring, experimental design and measurement system assessment. He is a Fellow of both the American Society for Quality and the American Statistical Association.

How to cite this article: Hamada MS, Casleton EM, Osthus D, Weaver BP, Steiner SH. Analyzing count data with measurement error. *Qual Reliab Eng Int*. 2022;1–11. https://doi.org/10.1002/qre.3078