Statistics in the Twenty-First Century: Special Volume
In Honour of Distinguished Professor Dr. Mir Masoom Ali
On the Occasion of his 75th Birthday Anniversary
PJSOR, Vol. 8, No. 3, pages 381-391, July 2012

Inference for Noisy Samples

Ibrahim. A. Ahmad
Department of Statistics
Oklahoma State University
Stillwater, Oklahoma 74078
iahmad@okstate.edu

N. Herawati
Department of Mathematics
University of Lampung
Bandar Lampung, Indonesia 35145

Abstract

In the current work, some well-known inference procedures including testing and estimation are adjusted to accommodate noisy data that lead to nonidentically distributed sample. The main two cases addressed are the Poisson and the normal distributions. Both one and two sample cases are addressed. Other cases including the exponential and the Pareto distributions are briefly mentioned. In the Poisson case, the situation when the sample size is random is mentioned.

Keywords: Noisy samples, Hypothesis testing and estimation, Poisson, Normal, Exponential, and Pareto.

1.Introduction

Statistical inference with its estimation and hypothesis testing routes is usually based on the concept of “random samples”.This means that the data framework is assumed to be “independent identically distributed” (iid) random variables. Although this setting is becoming less and less practical, it is still the way inference is presented to statistics students via textbooks even at the graduate level, see Rohatgi and Saleh (2001) and Casella and Berger (2002). When the iid structure is violated, there are no guarantees that standard accepted procedures will continue to hold and often modifications result in only partial and or weaker solutions. This need not be the case in some noisy sampling schemes that are now in use. In the current work, it is shown that for some non-iid sampling schemes in the Poisson and normal cases, standard procedures can be modified keeping the main characteristics of the estimates or test statistics intact. The sampling we discuss here includes the now popular rate sampling, cf. Thode (1997) and Ng and Tang (2005) for Poisson distribution and Moser et al. (1989), and Sprott and Farewell (1993) for the normal distribution, among others.

Rate sampling discussed here means that we have either a Poisson or Gaussian process and that we observe the sum of readings during adjacent time slots of lengths $C_1,...,C_n$, where $C_i > 0$, $i = 1,...,n$. Thus in the Poisson case, the observations are non-iid random variables X_i such that $E(X_i) = C_i\lambda$, $i = 1,...,n$. While in Gaussian case, we have $E(X_i) = C_i\mu$ and $V(X_i) = C_i^2\sigma^2$ or in the homogeneous variance case is just σ^2.

Variations of the rate sampling in the Poisson case have been researched in the literature including work by Detre and White (1970), Sichel (1973), Thode (1997), Krishnamoorthy and Thomsons (2004), and Ng and Tang (2005). Rate sampling in the normal case with homogenous variance is known as the regression through the origin case. The celebrated Behrens-Fisher problem is a two-sample variation and solutions for several different forms of it were discussed by Bernard (1982), Moser et al. (1989), Sprott and Farewell (1993), and Moser and Stevens (1992).

2. Poisson Inference

Let $X_1,...,X_n$ denote a random sample from Poisson distributions such that $X_i \sim P(\lambda_i)$ with $\lambda_i = C_i\lambda$, $i = 1,...,n$, where $C_i \geq 0$ is a known constant. Our task is to do inference about λ. The usual iid case falls at $C_1 = ... = C_n = 1$.

The likelihood function is:

$$L(\underline{X}, \underline{C}\lambda) = e^{-\lambda\sum C_i} \frac{\prod C_i^{X_i} \lambda^{\sum X_i}}{\prod X_i!}. \tag{1}$$

Clearly one obtains the maximum likelihood estimate (MLE) of λ to be

$$\hat{\lambda}_{MLE} = \hat{\lambda}_1 = \frac{\sum X_i}{\sum C_i}.$$

This estimate is complete and sufficient for λ as well as unbiased, so it is the uniformly minimum variance unbiased estimate (UMVUE) of λ. This estimate suggests the following class of unbiased estimates for λ:

$$\hat{\lambda}_{r+1} = \frac{\sum C_i^r X_i}{\sum C_i^{r+1}}, r \geq 0 \text{ is an integer.}$$

Now clearly,

$$V(\hat{\lambda}_{r+1}) = \lambda \frac{\sum C_i^{2r+1}}{(\sum C_i^{r+1})^2}. \tag{2}$$

But since $\hat{\lambda}_1$ is the UMVUE, we have

$$V(\hat{\lambda}_1) = \frac{\lambda}{\sum C_i} \leq V(\hat{\lambda}_{r+1}) = \lambda \frac{\sum C_i^{2r+1}}{(\sum C_i^{r+1})^2}. \tag{3}$$

Hence as a by product we have a simple proof of the inequality:

$$(\sum C_i^{r+1})^2 \leq (\sum C_i)(\sum C_i^{2r+1}), \tag{4}$$

where $C_i \geq 0$, i=1,..., n. Note also that in the Poisson case $E(X_i) = V(X_i), i = 1, \cdots\cdots, n$. Thus we might try to find an unbiased estimate of λ based on the sample variance. Precisely we have:

THEOREM(1): The following estimate is an unbiased estimate of λ :

$$\hat{\hat{\lambda}} = \frac{\sum C_i}{(\sum C_i)^2 - \sum C_i^2} \sum_i (X_i - C_i \hat{\lambda}_1)^2. \tag{5}$$

PROOF:Note that

$$E\sum_i (X_i - C_i\hat{\lambda}_1)^2 = \sum_i EX_i^2 - 2\sum_i C_i E\hat{\lambda}_1 X_i + \sum_i C_i^2 E(\hat{\lambda}_1^2). \tag{6}$$

But

$$EX_i^2 = C_i\lambda + C_i^2\lambda,$$

$$EX_i\hat{\lambda}_1 = \frac{1}{\Sigma C_i}\left\{C_i\lambda + C_i^2\lambda^2 + C_i\lambda^2 \sum_{j\neq i} C_j\right\} \text{ and}$$

$$E\hat{\lambda}_1^2 = V(\hat{\lambda}_1) + [E(\hat{\lambda}_1)]^2 = \frac{\lambda}{\Sigma C_i} + \lambda^2.$$

Summing and substituting into (6) yield that

$$E\sum_i (X_i - C_i\hat{\lambda}_1)^2 = \frac{(\Sigma C_i)^2 - \Sigma C_i^2}{\Sigma C_i}.$$

The result now follows.

Let us briefly discuss the relations between $\hat{\lambda}_1 = \hat{\lambda}$ and $\hat{\hat{\lambda}}$. Since $\hat{\lambda}$ is the UMVUE of λ, and if $E(\hat{\hat{\lambda}} \mid \hat{\lambda}) = \psi(\hat{\lambda})$, then ψ must be the identity function, i.e. $\psi(\hat{\lambda}) = \hat{\lambda}$. Hence

$$Cov(\hat{\lambda}, \hat{\hat{\lambda}}) = E(\hat{\lambda}\hat{\hat{\lambda}}) - \lambda^2 = E\left\{\hat{\lambda}E(\hat{\hat{\lambda}} \mid \hat{\lambda})\right\} - \lambda^2 = V(\hat{\lambda}). \tag{7}$$

Hence $\hat{\lambda}$ and $\hat{\hat{\lambda}}$ are positively correlated with correlation coefficient equal to $\rho(\hat{\lambda}, \hat{\hat{\lambda}}) = \{\frac{V(\hat{\lambda})}{V(\hat{\hat{\lambda}})}\}^{1/2}$. In the special case $C_1 = \cdots\cdots = C_n = 1, V(\hat{\hat{\lambda}}) = \lambda / n + 2\lambda^2 /(n-1)$, thus

$$\rho\left(\hat{\lambda}, \hat{\hat{\lambda}}\right) = \frac{1}{\sqrt{1 + 2\lambda n/(n-1)}} \to \frac{1}{\sqrt{1+2\lambda}} \text{ as } n \to \infty. \tag{8}$$

We may combine $\hat{\lambda}$ and $\hat{\hat{\lambda}}$ to obtain yet another unbiased estimate of λ.

Note that $E(\hat{\hat{\lambda}} / \hat{\lambda}) = E\left\{\frac{1}{\hat{\lambda}} E(\hat{\hat{\lambda}} \mid \hat{\lambda})\right\} = 1.$ Thus

$$\left\{\frac{\Sigma C_i}{(\Sigma C_i)^2 - \Sigma C_i^2}\right\} E\left\{\frac{\sum_i (X_i - C_i\hat{\lambda})^2}{\hat{\lambda}}\right\} = 1,$$

and this simplifies to

$$\lambda \sum C_i^2 = E\left\{\frac{\sum X_i^2}{\sum X_i}\right\} - \frac{(\sum C_i)^2 - \sum C_i^2}{\sum C_i}.$$

Hence

$$\lambda^{*}=\frac{1}{\sum C_i^2}\left\{\frac{\left(\sum C_i\right)\left(\sum X_i^2\right)}{\sum X_i}-\frac{\left(\sum C_i\right)^2-\sum C_i^2}{\left(\sum C_i\right)}\right\} \tag{9}$$

is an unbiased estimate of λ .

Sometime in sampling from Poisson distribution we do inverse sampling where the sample size is random. Let N be an integer-valued random variable not necessarily independent of X_i 's. We then estimate λ by

$$\hat{\lambda}_{r,N}=\frac{\sum_{i=1}^{N}C_i^r X_i^r}{\sum_{i=1}^{N}C_i^{r+1}}, r\geq 0, \tag{10}$$

and

$$\hat{\hat{\lambda}}_N=\frac{\sum_{i=1}^{N}C_i}{(\sum C_i)^2-\sum C_i^2}\sum_{i=1}^{N}(X_i-C_i\hat{\lambda}_{0,N})^2. \tag{11}$$

Clearly, $E\left(\hat{\lambda}_{r,N}\right)=E\left\{E\left(\lambda_{r,N}\mid N\right)\right\}=\lambda,$ and

$$V\left(\lambda_{r,N}\right)=E\left\{V\left(\hat{\lambda}_{r,N}\mid N\right)\right\}+V\left\{E\left(\hat{\lambda}_{r,N}\mid N\right)\right\}=\lambda E\left\{\frac{\sum_1^N C_i^{2r+1}}{\left(\sum_1^N C_i^{r+1}\right)^2}\right\}. \tag{12}$$

In the usual case when $C_1=\cdots=C_n=1$, we get that

$$V\left(\lambda_{r,N}\right)=\lambda E\{1/N\}.$$

On the other hand,

$$V(\hat{\hat{\lambda}}_N)=E\left\{V(\hat{\hat{\lambda}}_N\mid N)\right\}=\lambda\left[E\left(\tfrac{1}{N}\right)+2\lambda E\left(\tfrac{1}{N-1}\right)\right]. \tag{13}$$

Thus the correlation coefficient between $\hat{\lambda}_N$ and $\hat{\hat{\lambda}}_N$ is

$$\rho\left(\hat{\lambda}_N,\hat{\hat{\lambda}}_N\right)=\frac{1}{\sqrt{1+2\lambda\frac{E(\frac{1}{N})}{E(\frac{1}{N-1})}}}\rightarrow\frac{1}{\sqrt{1+2K\lambda}},$$

provided that

$$E(\tfrac{1}{N})/E(\tfrac{1}{N-1})\rightarrow K>0.$$

Finally, we can easily deduce the following unbiased estimate:

$$\lambda_N^{*}=\left(\frac{\sum_1^N X_i^2}{\sum_1^N X_i}\right)-\left(\frac{N-1}{N}\right). \tag{14}$$

To test $H_0:\lambda=\lambda_0$ vs. $H_1:\lambda>\lambda_0$ or $\lambda<\lambda_0$ or $\lambda\neq\lambda_0$, set any of the following test statistics that are all asymptotically normal:

$$Z_1=\sqrt{\Sigma C_i}\cdot\frac{\hat{\lambda}-\lambda_0}{\sqrt{\hat{\lambda}}} \text{ or } Z_2=\sqrt{\Sigma C_i}\cdot\frac{\hat{\lambda}-\lambda_0}{\sqrt{\hat{\hat{\lambda}}}}. \qquad (15)$$

Next, let us consider the two-sample case: let $X_i\sim P(C_i\lambda_1), i=1,\cdots,m$ and $Y_j\sim P(D_j\lambda_2), j=1,\cdots,n$. Assume the X_i's and the Y_j's are independent. We want to test $H_0:\lambda_1=R_0\lambda_2$, where R_0 is a known constant. Again, $\hat{\lambda}_1=\sum X_i/\sum C_i$ and $\hat{\lambda}_2=\sum Y_i/\sum D_i$. Further, under H_0, we have the likelihood function (with $\lambda_2=\lambda$) given by:

$$L(\underline{X},\underline{Y},\underline{C},\underline{D},R_0,\lambda)=\Pi\frac{C_i^{X_i}}{X_i}\Pi\frac{D_i^{Y_i}}{Y_i}e^{-\lambda(R_0\Sigma C_i+\Sigma D_j^2)}\lambda^{\Sigma X_i+\Sigma Y_j}R_0^{\Sigma X_i}. \qquad (16)$$

Hence (16) leads to the null estimate of λ:

$$\hat{\lambda}_0=\frac{\sum X_i+\sum Y_j}{R_0\sum C_i+\sum D_j}. \qquad (17)$$

Set the test statistic:

$$Z_3=\frac{R_0\hat{\lambda}_1-\hat{\lambda}_2}{\sqrt{\hat{\lambda}_0(\frac{R_0}{\sum C_i}+\frac{1}{\sum D_j})}}=\frac{R_0\frac{\sum X_i}{\sum C_i}-\frac{\sum Y_j}{\sum D_j}}{\sqrt{(\frac{\sum X_i+\sum Y_j}{R_0\sum C_i+\sum D_j})(\frac{R_0}{\sum C_i}+\frac{1}{\sum D_j})}}. \qquad (18)$$

Reject H_0 when $|Z_3|>Z_\alpha$ or $Z_{\alpha/2}$ according to the alternative. If we would like an unbiased variance null estimate for λ we use the following:

THEOREM (2): Under H_0, the following is an unbiased estimate for the common parameter λ_0

$$\hat{\hat{\lambda}}_0=[\frac{R_0\sum C_i+\sum D_j}{\sqrt{(R_0\sum C_i+\sum D_j)^2-(R_0^2\sum C_i^2+\sum D_j^2)}}][\sum_i(X_i-C_i\hat{\lambda})^2+\sum_j(Y_j-D_j\hat{\lambda})^2], \qquad (19)$$

where $\hat{\lambda}_0$ is as given in (17).

PROOF: Let

$$\begin{aligned}\hat{\eta}&=\sum_i(X_i-R_0C_i\hat{\lambda})^2+\sum_j(Y_j-D_j\hat{\lambda})^2\\&=\sum X_i^2+\sum Y_j^2++\hat{\lambda}^2(R_0^2\sum C_i^2+\sum D_j^2)-2\hat{\lambda}(R_0\sum C_iX_i+\sum D_jY_j)..\end{aligned} \qquad (20)$$

The result follows by noticing that:

$$E(\sum X_i^2 + \sum Y_j^2) = R_0\lambda\sum C_i + R_0^2\lambda^2\sum C_i^2 + \lambda\sum D_j + \lambda^2\sum D_j^2, \tag{21}$$

$$R_0 E(\sum C_i\lambda\hat{\lambda}) = \frac{R_0^2}{R_0\sum C_i + \sum D_j}(\lambda\sum C_i^2 + \lambda^2 R_0\sum C_i\sum C_i^2 + \lambda^2 R_0\sum C_i^2\sum D_j), \tag{22}$$

$$E(\sum D_j Y_j\hat{\lambda}) = \frac{1}{R_0\sum C_i + \sum D_j}(\lambda\sum D_j^2 + \lambda^2\sum D_j\sum D_j^2 + \lambda^2\sum C_i\sum D_j^2) \tag{23}$$

and

$$E(\hat{\lambda}^2) = V(\hat{\lambda}) + (E(\hat{\lambda}))^2 = \frac{\lambda}{R_0\sum C_i + \sum D_j} + \lambda^2. \tag{24}$$

The result follows from (21) to (24).

A test statistic based on $\hat{\hat{\lambda}}_0$ is:

$$Z_4 = \frac{R_0\hat{\lambda}_1 - \hat{\lambda}_2}{\sqrt{\hat{\hat{\lambda}}_0(\frac{R_0}{\sum C_i} + \frac{1}{\sum D_j})}}. \tag{25}$$

Reject H_0 if $|Z_4| > Z_\alpha$ or $Z_{\alpha/2}$ according to the alternative.

3. Normal Inference

Let $X_1, X_2, \cdots, X_n$ denote a random sample from normal distribution such that $X_i \sim N(C_i\mu, \sigma^2)$, $i = 1, \cdots, n$. Then the likelihood function is given by

$$L(\underline{X}, \underline{C}, \mu, \sigma^2) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}\exp[\frac{-1}{2\sigma^2}\sum_i (X_i - C_i\mu)^2]. \tag{26}$$

Hence the MLE of μ and σ^2 are respectively:

$$\hat{\mu}_{\underline{C}} = \frac{\sum C_i X_i}{\sum C_i^2} \text{ and } \hat{\sigma}^2 = \frac{1}{n}\sum_i (X_i - C_i\hat{\mu})^2. \tag{27}$$

Clearly, $\hat{\mu}$ is sufficient, complete and unbiased for μ so it is the UMVUE, while, $E\hat{\sigma}^2 = (n-1)/n\sigma^2$, thus the UMVUE for σ^2 is $S_{\underline{C}}^2 = 1/(n-1)\sum_i (X_i - C_i\hat{\mu})^2$.

Furthermore, we can verify that $\hat{\mu}$ and $S_{\underline{C}}^2$ are independent and $\hat{\mu}_{\underline{C}}$ is normal with mean μ and variance $\sigma^2/\sum C_i^2$ and $(n-1)S_{\underline{C}}^2/\sigma^2$ is $\chi^2(n-1)$ as expected, since

$$\sum (X_i - C_i\hat{\mu})^2 = \sum (X_i - C_i\mu)^2 - n(\hat{\mu}_{\underline{C}} - \mu)^2.$$

To test $H_0 : \mu = \mu_0$, we use the modified $t(n-1)$ statistic $T = \sqrt{\Sigma C_i^2}(\hat{\mu} - \mu_0)/S_{\underline{C}}$.

Let us turn our attention to a general setting of the two sample case: let $X_i \sim N(C_i\mu_1, \sigma^2)$, $i = 1, \cdots, m$ and $Y_i \sim N(D_j\mu_2, \gamma\sigma^2)$, $j = 1, \cdots, n$, where $C_1, \cdots, C_m, D_1, \cdots, D_n$ and γ are known constant. We want to test $H_0 : \mu_1 = \alpha\mu_2 + \beta$ for given values α and β. Note that the above setting include most known cases where: (*i*) $\alpha = 1, \beta = 0$ is testing $H_0 : \mu_1 = \mu_2$, (*ii*) $\alpha = 1, \beta$ given in testing $H_0 : \mu_1 - \mu_2 = \beta$, (*iii*) $\beta = 0$ and given α is testing $H_0 : \mu_1 = \alpha\mu_2$. As shown above the UMVUE's for μ_1 and μ_2 are, respectively, $\hat{\mu}_1 = \frac{\sum C_i X_i}{\sum C_i^2}$ and $\hat{\mu}_2 = \frac{\sum D_j Y_j}{\sum D_j^2}$. Now, we need to set an unbiased null variance estimate which is independent of $\hat{\mu}_1$ and $\hat{\mu}_2$. This is done as the following:

THEOREM (3): Under $H_0 : \mu_1 = \alpha\mu_2 + \beta$ and with $\mu_2 = \mu$, the MLE estimates of μ and σ^2 (adjusted to be unbiased) are:

$$\hat{\mu}_0 = \frac{\alpha \sum C_i X_i + \frac{1}{\gamma} \sum D_j Y_j - \alpha\beta \sum C_i^2}{\alpha^2 \sum C_i^2 + \frac{1}{\gamma} \sum D_j^2}, \tag{28}$$

and

$$\hat{\sigma}_0^2 = \frac{1}{m+n-2} [\sum_i (X_i - \alpha C_i \hat{\mu}_0 - C_i \beta)^2 + \frac{1}{\gamma} \sum_j (Y_j - \hat{\mu}_0 D_j)^2]. \tag{29}$$

PROOF:

Under H_0, the likelihood function of the two samples is given by:

$$L(\underline{X}, \underline{Y}, \underline{C}, \underline{D}, \alpha, \beta, \gamma; \mu, \sigma^2)$$

$$= (\frac{1}{\sqrt{2\pi}})^{m+n} (\sigma^2)^{\frac{m+n}{2}} \gamma^n \exp\{\frac{-1}{2\sigma^2} [\sum_i (X_i - \alpha C_i \mu - C_i \beta)^2 + \frac{1}{\gamma} \sum_j (Y_j - D_j \mu)]\} \tag{30}$$

Taking the log of L and differentiating with respect to μ results in (28). Then differentiating with respect to σ^2 results in $\sigma_0^{*2} = \frac{m+n}{m+n-1} \hat{\sigma}_0^2$. Correcting the bias comes through evaluating the following:

$$E\hat{\eta} = E[\sum_i (X_i - \alpha C_i \hat{\mu}_0 - C_i \beta)^2 + \frac{1}{\gamma} \sum_j (Y_j - D_j \frac{1}{\gamma} \sum_j (Y_j - D_j \hat{\mu}_0)^2]$$

$$= \sum_i V(X_i) + \frac{1}{\gamma} \sum_j V(Y_j) - 2V(\hat{\mu}_0)(\alpha^2 \sum C_i^2 + \frac{1}{\gamma} \sum_j D_j^2)$$

$$= \sigma^2 (m+n). \tag{31}$$

But $V(\hat{\mu}_0) = \dfrac{\sigma^2}{(\alpha^2 \sum C_i^2 + \frac{1}{\gamma}\sum_j D_j^2)}$. The result now follows.

Hence, set the test statistic

$$T = \frac{\hat{\mu}_1 - \alpha\hat{\mu}_2 - \beta}{\hat{\sigma}_0 \sqrt{\dfrac{1}{\sum C_i^2} + \dfrac{\alpha^2}{\gamma \sum_j D_j^2}}} \sim t(m+n-2). \tag{32}$$

4. Other Examples and Final Remarks

(A) Other examples:

Extending the MLE's and the hypotheses testing presented earlier for the non-iid samples from the Poisson and the normal distributions can be used in various other cases. We illustrate this by two further examples.

Example (1): Let $X_i \sim \exp(C_i\mu, \sigma)$ with p.d.f. $f_i(x) = \sigma^{-1}\exp[-(\frac{x - C_i\mu}{\sigma})]$, where $x \geq \{\max_{1\leq i\leq n} C_i\}\mu$ and $\sigma > 0$, $i = 1, 2, ..., \mu$. The likelihood function is:

$$L(\underline{X}, \underline{C}; \mu, \sigma) = \prod_{i=1}^{n} \sigma^{-1}\exp[-(1/\sigma)(X_i - C_i\mu)], \min_i(\frac{X_i}{C_i}) \geq \mu. \tag{33}$$

Hence the MLE of μ is $\hat{\mu} = \min_{1\leq i\leq n}(X_i/C_i)$ and of σ is $\hat{\sigma} = (1/n)\sum_i (X_i - C_i\hat{\mu})$.But clearly, $E(\hat{\mu}) = \mu + \dfrac{\sigma}{\sum C_i}$ and $E(\hat{\sigma}) = \dfrac{n-1}{n}\sigma.$

This leads to the UMVUE's of μ and σ given by:

$$\hat{\mu}_{UB} = \frac{n}{n-1}\left\{\min_{1\leq i\leq n}\left(\frac{X_i}{C_i}\right) - \hat{\mu}\sum C_i\right\} \tag{34}$$

and

$$\hat{\sigma}_{UB} = \frac{1}{n-1}\sum_i (X_i - C_i\hat{\mu}). \tag{35}$$

When $C_1 = C_2 = \cdots\cdots = C_n = 1$, we get the known unbiased estimate of μ, $\hat{\mu} = \dfrac{X_{(1)} - \overline{X}/n}{n-1}$.

Example (2): Let $X_i \sim f_i(x) = \dfrac{1}{\theta^{C_i}}, 0 < x \leq \theta^{C_i}$, i =1,2,…,n.

Thus the likelihood function is

$$L(\underline{X},\underline{C};\theta)=\theta^{-\sum C_i},\theta\geq\max_{1\leq i\leq n}(X_i)^{\frac{1}{C_i}}. \tag{36}$$

Hence the MLE of θ is

$$\hat{\theta}=\max_{1\leq i\leq n}(X_i^{\frac{1}{C_i}}). \tag{37}$$

To find the UMVUE of θ we need to adjust $\hat{\theta}$ for bias. Note that

$$E\left(\hat{\theta}\right)=\int_0^\theta P(\hat{\theta}>u)du=\theta-\int_0^\theta \pi_i F_i(x)dx=\theta\frac{\sum C_i}{\sum C_i+1}. \tag{38}$$

Hence $\hat{\theta}_{UB}=\left(\frac{\sum C_i}{\sum C_i+1}\right)\max_{1\leq i\leq n}(X_i^{\frac{1}{C_i}})$ is the UMVUE of θ.

(B) Some Final Remarks:

(i) In the normal two-sample case, there is a major outstanding problem known as the "Behrens-Fisher" which is that when $\sigma_1\neq\sigma_2$ there is no estimate for $\left(\sigma_1^2+\alpha^2\sigma_2^2\right)$ that can lead to a t-statistic, cf, Rohatgi and Saleh (2001). There are many suggestions of partial solutions. In this context, we offered one when we can assume $\sigma_1^2=\gamma\sigma_2^2$, see Section 2. Intermediate, approximate or partial solutions have been offered in the literature. For example, when one σ is known, say σ_1, or in the general case of $\sigma_1\neq\sigma_2$, both unknown, Satterthwaite (1941) and (1946) proposed a method that provides an approximate χ^2 for $\gamma\left(\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}\right)^{-1}\left(\frac{S_1^2}{m}+\frac{S_2^2}{n}\right)$ with $\gamma=\left(\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}\right)^2\left\{\frac{\left(\sigma_1^2/m\right)^2}{m-1}+\frac{\left(\sigma_2^2/n\right)^2}{n-1}\right\}^{-1}$. Hence γ could be estimated by replacing σ_1^2 and σ_2^2 by S_1^2 and S_2^2, respectively, see also Ames and Webster (1991) and Moser and Stevens (1992). The case when σ_1 is given was discussed by Maity and Sherman (2006). The value of γ is obtained via matching $(\gamma^*)(\sigma_1^2/m+\sigma_2^2/n)V(S_1^2/m+S_2^2/n)=2\gamma^*$. This gives $\gamma^*=(\sigma_1^2/m+\sigma_2^2/n)^2\{\sigma_2^4/[n(n-1)]\}$. Then one replaces σ_2^2 with its estimate S_2^2. The above method is based on matching the variance of the proposed estimate of $\left(\sigma_1^2+\alpha^2\sigma_2^2\right)$(standardized) to the variance of a Chi-square $\left(m+n-2\right)$ and solving for the induced parameter γ and then estimate it. Needless to say matching variance is not enough to guarantee matching of distributions. In a subsequent work we will provide a new method based on matching the distributions.

(ii) Note that the essence of the Behrens-Fisher problem for testing $H_0 : \mu_2 = \alpha\mu_1 + \beta$, α, β known constants is that the estimate of $V(\overline{Y} - \alpha\overline{Y} - \beta) = \frac{\sigma_1^2}{n} + \alpha^2 \frac{\sigma_1^2}{n}$ which is a weighted sum of the $(m-1)$ and $(n-1)$ Chi-squares with unknown weights $K_1(m,n) = \frac{\sigma_1^{\ 2}}{m(m-1)}\left(\frac{\sigma_1^{\ 2}}{m} + \alpha^2 \frac{\sigma_2^{\ 2}}{n}\right)^{-1}$ and $K_2(m,n) = \frac{\sigma_2^{\ 2}}{n(n-1)}\left(\frac{\sigma_1^{\ 2}}{m} + \alpha^2 \frac{\sigma_2^{\ 2}}{n}\right)^{-1}$. When we replace $K_1(m,n)$, $i = 1,2$ by their estimates based on S_1^2 and S_2^2 we losethe independence and the problem becomes intractable. We shall provide in the near future some plausible solutions based on what the weighted sum of Chi-squares looks like.

(iii) There are situations where even for the model we discussed above, namely that the parameters are known to beproportionate, there may not be an explicit solution. However, in some of these cases approximate solutions may be possible. Let us illustrate this via the example of nonidentical Bernoulli trials: let $X_1,...,X_n$ be n independent Bernoulli rv's such that $P(X_i = 1) = p_i = C_i p,\ i = 1,...,n$. As before assume $C_i \geq 0$, $i = 1,...,n$ to be known. Thus, the likelihood function is:

$$L(\underline{X}, \underline{C}, p) = \prod_i C_i^{X_i} p^{\sum X_i} \prod_i (1 - C_i p)^{1-X_i} \tag{39}$$

Taking the log of $L(\underline{X}, \underline{C}, p)$ and differentiating with respect to p we get

$$\sum_i X_i - \sum_i \left(\frac{C_i p}{1 - C_i p}\right)(1 - X_i) = 0 \tag{40}$$

It is not difficult to show that the MLE of p exists and is unique but it is not explicit. If we need an explicit solution, we can use the approximation $\frac{C_i p}{1-C_i p} \approx C_i p, i = 1,...,n$, thus solve

$$\sum_i X_i - p\sum_i C_i (1 - X_i) = 0,$$

giving an approximate estimate $\hat{p} = \frac{\sum_i X_i}{\sum_i C_i (1 - X_i)}$. Using the well known formula $E(\frac{X}{Y}) \approx \frac{E(X)}{E(Y)}$, we get that

$$E(\hat{p}) \approx p\{1 + \frac{p\sum_i C_i^{\ 2}}{\sum_i C_i + p\sum_i C_i^{\ 2}}\}. \tag{41}$$

Thus if $\sum C_i^{\ 2} / \sum C_i \to K \geq 0$ then $E(\hat{p}) \approx p\{1 + \frac{pK}{1+pK}\}$. Hence $\hat{p}$ is asymptotically unbiased if K=0 while it is positively biased if K>0 in which case, the estimate

$$\hat{\hat{p}} = \hat{p}\{1 + \frac{\hat{p}K}{1 + \hat{p}K}\}^{-1}, \tag{42}$$

is asymptotically unbiased. Finally, using the well known formula

$$V(\frac{X}{Y}) \approx [\frac{E(X)}{E(Y)}]^2 \{\frac{V(X)}{[E(X)]^2} + \frac{V(Y)}{[E(Y)]^2} - 2\frac{\text{cov}(X_1 Y)}{E(X)E(Y)}\}. \quad (43)$$

We get that

$$V(\hat{p}) \approx p^2 [\frac{\sum C_i}{\sum C_i (1 - C_i p)}]^2 \{\frac{\sum C_i (1 - C_i p)}{p(\sum C_i)^2} + \frac{p \sum C_i^{\,3} (1 - C_i p)}{[\sum C_i (1 - C_i p)]^2} + \frac{2}{\sum C_i}\}. \quad (44)$$

References

1. Ames, M. H. and Webster, J. T. (1991), "On Estimating Approximate Degrees of Freedom," *The American Statistician*, 45, 45-50.
2. Bernard, G. A. (1982), "A New Approach to the Behrens-Fisher Problem," *Utilitas Mathematica*, 218, 261-277.
3. Chamberlin, S. R., and Sprott, D. A. (1989), "The Estimation of a Location Parameter when the Scale Parameter is Confined to a Finite Range: The Notion of Generalized Ancillary Statistic," *Biometrika*, 76, 609-612.
4. Casella, G., and Berger, R. L. (2002), "Statistical Inference," 2nd Ed., Duxbury, Pacific Grove, CA.
5. Detre, K., and White, C. (1970), "The Comparison for Two Poisson Distributed Observations," *Biometrics*, 26, 851-854.
6. Krishnamoorthy, K., and Thomson, J. (2004). "A more Powerful Test for comparing Two Poisson Means," *Journal of Statistical Planning and Inference*, 119, 23-35.
7. Maity, A., and Sherman, M. (2006). "The Two-Sample T Test with One Variance known," *The American Statistician*, 60, 163-166.
8. Moser, B. K., and Stevens, G. R. (1992). "Homogeneity of Variance in the Two Sample Means Test," *The American Statistician*, 46, 19-21.
9. Moser, B. K., Stevens, G. R., and Walts, C. L. (1989). "The Two Sample t Test Versus Satterthwaite's Approximate F-Test," *Communications in Statistics-Theory and Methods*, 18, 3963-3975.
10. Ng, H. K. T., and Tang, M. L. (2005). "Testing the Equality of Two Poisson Means Using the Rate Ratio," *Statistics in Medicine*, 24, 955-965.
11. Rohatgi, V. K., and Saleh, Md. K. E. (2001). "An Introduction to Probability and Statistics," 2nd Ed., Wiley&Sons, New York, NY.
12. Satterthwaite, F. E. (1941). "Synthesis of Variance," *Psychometrika*, 6, 309-316.
13. Satterthwaite, F. E. (1946). "An Approximatie Distribution of Estimates of Variance Component," *Biometric Bulletin*, 6, 110-114.
14. Sichel, H. S. (1973). "On a Significance Test for Two Poisson Variables," *Applied Statistics*, 22, 50-58.
15. Sprott, D. A., and Farewell, V. T. (1993). "The Differences Between Two Normal Means," *The American Statisticians*, 47, 126-128.
16. Thode, H. C. (1997). "Power and Sample Size Requirements for Tests of Differences Between Two Poisson Rates," *The Statistician*, 46, 227-230.