Bayesian Estimation of Functions of Binomial Rates

L. Raubenheimer

Department of Statistics, Rhodes University, Grahamstown, South Africa

e-mail: *L.Raubenheimer@ru.ac.za*

and

A.J. van der Merwe

Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

---

Key words: Bayesian intervals, Coverage probabilities, Product of Binomial proportions, Difference of Binomial proportions, Probability matching prior, Weighted Monte Carlo method.

Summary: In this paper the probability matching priors for the product of k Binomial parameters as well as for a linear combination of Binomial parameters are derived. In the case of two independently distributed Binomial variables, the Jeffreys', uniform and probability matching priors for the product of the parameters (ratios) are compared. The construction of Bayesian confidence intervals for the difference of two independent Binomial parameters is also discussed. This research is an extension of the work by Kim (2006) who derived a probability matching prior for the product of k independent Poisson rates.

---

1. Introduction

The Bayesian paradigm emerges as attractive in many types of statistical problems - especially in Binomial and Poisson problems but the choice of an appropriate non-informative prior distribution has been controversial. Common non-informative priors in multiparameter problems, such as Jeffreys' prior can have features that have an unexpectedly dramatic effect on the posterior distribution. Datta & Ghosh (1995) derived the differential equation which a prior must satisfy if the posterior probability of a one sided credibility interval for a parametric function and its frequentist probability agree up to $O\left(n^{-1}\right)$ where n is the sample size. They proved that the agreement between the posterior probability and the frequentist probability holds if and only if $\sum_{i=1}^{k} \frac{\partial}{\partial p_i}\left\{\eta_i\left(\underline{p}\right)\pi\left(\underline{p}\right)\right\} = 0$, where $\pi\left(\underline{p}\right)$ is the probability matching prior for $\underline{p}$, the vector of unknown parameters. Let $\nabla_t = \left[\begin{array}{ccc} \frac{\partial}{\partial p_1}t\left(\underline{p}\right) & \cdots & \frac{\partial}{\partial p_k}t\left(\underline{p}\right) \end{array}\right]'$, then $\eta\left(\underline{p}\right) = \frac{F^{-1}\left(\underline{p}\right)\nabla_t\left(\underline{p}\right)}{\sqrt{\nabla_t'\left(\underline{p}\right)F^{-1}\left(\underline{p}\right)\nabla_t\left(\underline{p}\right)}} = \left[\begin{array}{ccc} \eta_1\left(\underline{p}\right) & \cdots & \eta_k\left(\underline{p}\right) \end{array}\right]'$. Where $F^{-1}\left(\underline{p}\right)$ is the inverse of $F\left(\underline{p}\right)$, the

Fisher information matrix of $\underline{p}$. Reasons for using the probability matching prior is that it provides a method of constructing accurate frequentist intervals and it could also be useful for comparative purposes in Bayesian analysis. From Wolpert (2004), Berger states that frequentist reasoning will play an important role in finally obtaining good general priors for estimation and prediction. Some statisticians argue that frequency calculations are an important part of applied Bayesian statistics. (See Rubin, 1984)

In the next section the probability matching prior for the product of k Binomial rates is derived and in Section a weighted Monte Carlo simulation method is described to obtain Bayesian confidence intervals in the case of the probability matching prior. In Section simulation results and examples are given for $\psi_1 = p_1p_2$ and $\psi_2 = p_1p_2p_3$ and in Section the probability matching prior for $\theta = \sum_{i=1}^{k} \alpha_i p_i$, a linear combination of Binomial parameters is derived. Simulation results are discussed for $\theta = p_1 - p_2$ and the Jeffreys', uniform and probability matching priors are applied to a real problem.

2. Probability Matching Prior for the Product of Different Powers of *k* Binomial Parameters

The parameter $\psi = \prod_{i=1}^{k} p_i^{\alpha_i}$, the product of different powers of k Binomial parameters, appears in applications to system reliability. Assume for example systems that consist of two and three components in parallel, respectively, then the probabilities of system failure are $\psi_1 = p_1p_2$ and $\psi_2 = p_1p_2p_3$, the product of two and three Binomial proportions. Also assume that a system requires that at least one of each of two types of components must be employed and that these components in parallel are needed. The probability of failure of a system is then either $\psi_3 = p_1^2p_2$ or $\psi_4 = p_1p_2^2$ depending on whether the first or second component has been replicated, (Kim, 2006).

From a Bayesian perspective a prior is needed for the parameter ψ. As mentioned common non-informative priors in multiparameter problems such as Jeffreys' priors can have features that have an unexpectedly dramatic effect on the posterior distribution. It is for this reason that the probability matching prior for ψ will be derived in Theorem 2.1.

Also as mentioned a probability matching prior is a prior distribution under which the posterior probabilities match their coverage probabilities. The fact that the resulting Bayesian posterior intervals of level $1-\alpha$ are also good frequentist confidence intervals at the same level is a very desirable situation. See also Bayarri & Berger (2004) and Severini *et al.* (2002) for general discussion. By using the method of Datta & Ghosh (1995) the following theorem can be proved.

Theorem 2.1 Assume that $X_1, X_2, \ldots, X_k$ are independent Binomial random variables with $X_i \sim Bin(n_i, p_i)$ for $i = 1, 2, \ldots, k$. The probability matching prior for $\psi = \prod_{i=1}^{k} p_i^{\alpha_i}$, the product of different powers of k Binomial rates, is given by

$$\pi_s\left(\underline{p}\right) = \pi_s\left(p_1, p_2, \ldots, p_k\right) \propto \left\{\sum_{i=1}^{k} \frac{\alpha_i^2\left(1-p_i\right)}{p_i}\right\}^{\frac{1}{2}} \prod_{i=1}^{k}\left(1-p_i\right)^{-1} \tag{1}$$

Proof. The likelihood function is given by
$L\left(p_1, p_2 \ldots, p_k\right) = L\left(\underline{p}\right) = \prod_{i=1}^{k} \left(\begin{array}{c} n_i \\ x_i \end{array} \right) p_i^{x_i}\left(1-p_i\right)^{n_i-x_i}$.

The Fisher information matrix is well known and for $n_1 = n_2 = \ldots = n_k = n$, n can be ignored for all practical purposes (i.e. $n = 1$). The inverse of the Fisher information matrix is then given by

$$F^{-1}\left(\underline{p}\right) = diag\left[\begin{array}{cccc} p_1\left(1-p_1\right) & p_2\left(1-p_2\right) & \ldots & p_k\left(1-p_k\right) \end{array}\right].$$

We are interested in a probability matching prior for $t\left(\underline{p}\right) = \psi = \prod_{i=1}^{k} p_i^{\alpha_i}$, the product of different powers of k Binomial parameters.

Now

$$\begin{aligned} \nabla_t'\left(\underline{p}\right) &= \left[\begin{array}{cccc} \frac{\partial t\left(\underline{p}\right)}{\partial p_1} & \frac{\partial t\left(\underline{p}\right)}{\partial p_2} & \cdots & \frac{\partial t\left(\underline{p}\right)}{\partial p_k} \end{array}\right] \\ &= \left[\begin{array}{cccc} \frac{\alpha_1}{p_1} & \frac{\alpha_2}{p_2} & \cdots & \frac{\alpha_k}{p_k} \end{array}\right] \prod_{i=1}^{k} p_i^{\alpha_i}. \end{aligned}$$

Also
$\nabla_t'\left(\underline{p}\right) F^{-1}\left(\underline{p}\right) = \left[\begin{array}{cccc} \alpha_1\left(1-p_1\right) & \alpha_2\left(1-p_2\right) & \cdots & \alpha_k\left(1-p_k\right) \end{array}\right] \prod_{i=1}^{k} p_i^{\alpha_i}$
and
$\nabla_t'\left(\underline{p}\right) F^{-1}\left(\underline{p}\right) \nabla_t\left(\underline{p}\right) = \left(\prod_{i=1}^{k} p_i^{\alpha_i}\right)^2 \sum_{i=1}^{k} \frac{\alpha_i^2\left(1-p_i\right)}{p_i}$.
Define

$$\begin{aligned} \eta'\left(\underline{p}\right) &= \frac{\nabla_t'\left(\underline{p}\right) F^{-1}\left(\underline{p}\right)}{\sqrt{\nabla_t'\left(\underline{p}\right) F^{-1}\left(\underline{p}\right) \nabla_t\left(\underline{p}\right)}} \\ &= \left[\begin{array}{cccc} \eta_1\left(\underline{p}\right) & \eta_2\left(\underline{p}\right) & \cdots & \eta_k\left(\underline{p}\right) \end{array}\right] \end{aligned}$$

where $\eta_i\left(\underline{p}\right) = \frac{\alpha_i\left(1-p_i\right)}{\sqrt{\sum_{i=1}^{k} \frac{\alpha_i^2\left(1-p_i\right)}{p_i}}}$ $\qquad (i = 1, 2, \ldots, k)$.

The prior $\pi\left(\underline{p}\right)$ is a probability matching prior if and only if the differential equation $\sum_{i=1}^{k}\frac{\partial}{\partial p_i}\left\{\eta_i\left(\underline{p}\right)\pi\left(\underline{p}\right)\right\} = 0$ is satisfied.

The differential equation will be satisfied if $\pi\left(\underline{p}\right)$ is

$$\pi_s\left(\underline{p}\right) \propto \left\{\sum_{i=1}^{k}\frac{\alpha_i^2(1-p_i)}{p_i}\right\}^{\frac{1}{2}}\prod_{i=1}^{k}\left(\alpha_i\left(1-p_i\right)\right)^{-1}.$$

When $\alpha_i = 1$, the probability matching prior for $\psi = \prod_{i=1}^{k} p_i$, will be

$$\pi_s\left(\underline{p}\right) \propto \left\{\sum_{i=1}^{k}\frac{(1-p_i)}{p_i}\right\}^{\frac{1}{2}}\prod_{i=1}^{k}\left(1-p_i\right)^{-1}.$$

If $k = 1$ and $\alpha = 1$, equation 1 becomes the Jeffreys' prior. For $\alpha_i = 1$, $(i = 1,2,\ldots,k)$, the posterior distribution in the case of the probability matching prior is given by

$$\pi_s\left(\underline{p}|data\right) \propto \left\{\sum_{i=1}^{k}\frac{(1-p_i)}{p_i}\right\}^{\frac{1}{2}}\prod_{i=1}^{k}p_i^{x_i}\left(1-p_i\right)^{n_i-x_i-1} \tag{2}$$

for $0 \leq p_i \leq 1$. ■

Theorem 2.2 $\pi_s\left(\underline{p}|data\right)$ is a proper posterior distribution if $x_i < n_i$, for $(i = 1,2,\ldots,k)$.

Proof.

$$\begin{aligned}\sum_{i=1}^{k}\frac{1-p_i}{p_i} &= \frac{\sum_{i=1}^{k}\left\{\frac{\prod_{i=1}^{k}p_i}{p_i}-\prod_{i=1}^{k}p_i\right\}}{\prod_{i=1}^{k}p_i}\\ &< \frac{\sum_{i=1}^{k}\frac{\prod_{i=1}^{k}p_i}{p_i}}{\prod_{i=1}^{k}p_i} < \frac{k}{\prod_{i=1}^{k}p_i}\end{aligned}$$

Therefore $\sqrt{\sum_{i=1}^{k}\frac{1-p_i}{p_i}}\prod_{i=1}^{k}p_i^{x_i}\left(1-p_i\right)^{n_i-x_i-1} < \sqrt{k}\prod_{i=1}^{k}p_i^{x_i-\frac{1}{2}}\left(1-p_i\right)^{n_i-x_i-1}$.

Each $\int_0^1 p_i^{x_i-\frac{1}{2}}\left(1-p_i\right)^{n_i-x_i-1}dp_i = Beta\left(x_i+\frac{1}{2},n_i-x_i\right)$ converges if $x_i < n_i$ $(i = 1,\ldots,k)$. ■

The Jeffreys' prior on the other hand is proportional to the square root of the determinant of the Fisher information matrix and is given by

$$\pi_u\left(\underline{p}\right) \propto \left|F\left(\underline{p}\right)\right|^{\frac{1}{2}} = \left(\prod_{i=1}^{k} \frac{1}{p_i(1-p_i)}\right)^{\frac{1}{2}}.$$

The joint posterior is

$$\pi_u\left(\underline{p}\,|X_1,X_2,\ldots,X_k\right) \propto \left\{\prod_{i=1}^{k} B\left(x_i+\tfrac{1}{2};n_i-x_i+\tfrac{1}{2}\right)\right\}^{-1} \prod_{i=1}^{k} p_i^{x_i-\frac{1}{2}}\left(1-p_i\right)^{n_i-x_i-\frac{1}{2}} \qquad 0 \le p_i \le 1.$$

3. The Weighted Monte Carlo Method in the Case of $\psi = \prod_{i=1}^{k} p_i$ - Sampling - Importance Resampling

In this section a weighted Monte Carlo method is described which will be used for simulation from the posterior distribution in the case of the probability matching prior. This method is especially suitable for computing Bayesian confidence (credibility) intervals. It does not require knowing the closed form of the marginal posterior distribution of ψ, only the kernel of the posterior distribution of $\{p_1, p_2, \ldots, p_k\}$ is needed.

As mentioned by Chen & Shao (1999), Kim (2006), Smith & Gelfand (1992), Guttman & Menzefrieke (2003), Skare *et al.* (2003) and Li (2007) the weighted Monte Carlo (sampling - importance re-sampling (SIR)) algorithm aims at drawing a random sample from a target distribution π, by first drawing a sample from a proposal distribution q, and from this a smaller sample is drawn with sample probabilities proportional to the importance ratios π/q. For the algorithm to be efficient, it is important that q is a good approximation for π. This means that q should not have too light tails when compared to π. For further details see Skare *et al.* (2003). In the case of credibility intervals it is not even necessary to draw the smaller sample. The weights (sample probabilities) are however important.

If a uniform prior is put on $\underline{p}$, the posterior (proposal) distribution is

$$q\left(\underline{p}\,|data\right) \propto \left\{\prod_{i=1}^{k} B\left(x_i+1;n_i-x_i+1\right)\right\}^{-1} \prod_{i=1}^{k} p_i^{x_i}\left(1-p_i\right)^{n_i-x_i} \qquad 0 \le p_i \le 1.$$

In the case of the probability matching prior, the posterior (target) distribution is (see equation 2)

$$\pi_s\left(\underline{p}\,|data\right) \propto \left\{\sum_{i=1}^{k} \frac{(1-p_i)}{p_i}\right\}^{\frac{1}{2}} \left\{\prod_{i=1}^{k} \left(1-p_i\right)^{-1}\right\} \left\{\prod_{i=1}^{k} p_i^{x_i}\left(1-p_i\right)^{n_i-x_i}\right\} \qquad 0 \le p_i \le 1$$

The sample probabilities are therefore proportional to

$$\frac{\pi_s\left(\underline{p}\,|data\right)}{q\left(\underline{p}\,|data\right)} = \left\{\sum_{i=1}^{k} \frac{(1-p_i)}{p_i}\right\}^{\frac{1}{2}} \left\{\prod_{i=1}^{k} \left(1-p_i\right)^{-1}\right\} \qquad 0 \le p_i \le 1$$

and the normalized weights are

$$\omega_\ell = \frac{\left\{\sum_{i=1}^{k} \frac{\left(1-p_i^{(l)}\right)}{p_i^{(l)}}\right\}^{\frac{1}{2}} \left\{\prod_{i=1}^{k}\left(1-p_i^{(l)}\right)^{-1}\right\}}{\sum_{l=1}^{n}\left\{\sum_{i=1}^{k} \frac{\left(1-p_i^{(l)}\right)}{p_i^{(l)}}\right\}^{\frac{1}{2}} \left\{\prod_{i=1}^{k}\left(1-p_i^{(l)}\right)^{-1}\right\}} \qquad l = 1,2,\ldots,n.$$

A straightforward way of doing the weighted Monte Carlo (WMC) method was proposed by Chen & Shao (1999). Details of the Monte Carlo method are as follows:

Step 1

Obtain a Monte Carlo sample $\left\{\left(p_1^{(l)}, p_2^{(l)} \ldots, p_k^{(l)}\right); l = 1,2,\ldots,n\right\}$ from the proposal distribution $q\left(\underline{p}|data\right)$ and calculate $\psi^{(l)} = \prod_{i=1}^{k} p_i^{(l)}$ for $l = 1,2,\ldots,n$.

Step 2

Sort $\left\{\psi^{(l)}, (l = 1,2,\ldots,n)\right\}$ to obtain the ordered values $\psi^{[1]} \le \psi^{[2]} \le \cdots \le \psi^{[n]}$.

Step 3

Each simulated ψ value has an associated weight. Therefore compute the weighted function $\omega_{(l)}$ associated with the *lth* ordered $\psi^{[l]}$ value.

Step 4

Add the weights up from left to right (from the first on) until one gets $\sum_{l=1}^{n_1} \omega_{(l)} = \alpha/2$. Write down the corresponding $\psi^{[n_1]}$ value and denote it as $\psi_{(\alpha/2)}$. Add the weights up from right to left (from the last back) until one gets $\sum_{l=n_2}^{n} \omega_{(l)} = \alpha/2$. Write down the corresponding $\psi^{[n_2]}$ value and denote it as $\psi_{(1-\alpha/2)}$.

Step 5

The $100(1-\alpha)\%$ Bayesian confidence interval is:

$\left(\psi_{(\alpha/2)}, \psi_{(1-\alpha/2)}\right)$.

4. Examples

4.1 Reliability of Independent Parallel Components System Kim (2006)

Consider the following example for the probability of failure of independent parallel components system, using the observed data values from Harris (1971). One has to assume that the two systems consist of two and three components in parallel, respectively. The respective probabilities of system failure will then be $\psi_1 = p_1 p_2$ and $\psi_2 = p_1 p_2 p_3$. ψ_1 is the product of two binomial parameters, and ψ_2 is the product of three binomial parameters. When manufacturing the components system, one may be interested in an upper bound of the confidence interval on the system failure. The upper bound of the confidence interval, $\psi_{j,(1-\alpha)}$, for each system failure, which is given by $P\left(0 \le \psi_j \le \psi_{j,(1-\alpha)}\right) = 1-\alpha$ will be estimated for

$j = 1, 2$. The estimate of $\psi_{j,(1-\alpha)}$ is therefore the upper end point of a one-sided $(1-\alpha)\,100\%$ confidence interval for ψ_j. The methods used to obtain the upper limit of the confidence interval, are: The likelihood ratio method by Madansky (1965); Randomized limit based method by Harris (1971); Bayesian method by Kim (2006); The last two columns in Table 1 are obtained from the probability matching prior and Jeffreys' prior for the product of k Binomial rates. Kim approximated the Binomial distribution by the Poisson distribution and obtained a probability matching prior for $\tilde{\theta} = \prod_{i=1}^{k} \lambda_i$, the product of k Poisson rates. The prior is $p(\lambda_1, \lambda_2, \ldots \lambda_k) \propto \sqrt{\sum_{i=1}^{k} \lambda_i^{-1}}$. A simulated value for ψ is then obtained from the linear relationship between $\tilde{\theta}$ and ψ, namely $\psi = \tilde{\theta} / \prod_{i=1}^{k} n_i$. Comparisons between these five estimates are made in Table 1. The values for Madansky's and Harris's method are from Harris (1971) and the values for the Bayesian method are from Kim (2006).

The effectiveness of the comparisons between the five methods in Table 1 is rather restricted, since the five methods are all approximate and we do not have the exact confidence coefficient.

Table 1. Upper confidence limits for $\prod_{i=1}^{k} p_i$ with confidence coefficient $1-\alpha = 0.9$

Sample sizes n_1, n_2	**Observed** x_1, x_2	**Madansky's Method**	**Harris's Method**	**Bayesian Method**	**Probability Matching Prior**	**Jeffreys' Prior**
100, 100	3, 5	0.00433	0.00416	0.00406	0.00393	0.00355
100, 100	1, 4	0.00188	0.00184	0.00172	0.00167	0.00145
100, 100	2, 2	0.00168	0.00170	0.00157	0.00155	0.00131
150, 150	3, 3	0.00133	0.00128	0.00124	0.00120	0.00107

Sample sizes n_1, n_2, n_3	**Observed** x_1, x_2, x_3	**Madansky's Method**	**Harris's Method**	**Bayesian Method**	**Probability Matching Prior**	**Jeffreys' Prior**
100, 100, 100	1, 2, 1	0.000019	0.000027	0.000021	0.000021	0.000013
100, 100, 100	2, 3, 5	0.000133	0.000145	0.000132	0.000129	0.000102

As mentioned the last two columns are added to Table 2 of Kim (2006) and give $\psi_{(1-\alpha)}$ for the probability matching and Jeffreys' priors of $\psi_1 = p_1 p_2$ and $\psi_2 = p_1 p_2 p_3$. The values of $\psi_{(1-\alpha)}$ in the case of the probability matching prior compare well with those obtained by the other researcher while it seems that the Jeffreys' prior tends to some what under estimate the upper confidence limit.

4.2 Simulation Study Comparing Four Priors

In Tables 2 and 3 comparisons are made between the four priors:

1. $\pi_s\left(\underline{p}\right) \propto \left\{\sum_{i=1}^{k} \frac{(1-p_i)}{p_i}\right\}^{\frac{1}{2}} \prod_{i=1}^{k} (1-p_i)^{-1}$

2. $\pi_u\left(\underline{p}\right) \propto \left(\prod_{i=1}^{k} \frac{1}{p_i(1-p_i)}\right)^{\frac{1}{2}}$

3. $\pi_s\left(\underline{\lambda}\right) \propto \sqrt{\sum_{i=1}^{k} \lambda_i^{-1}}$

4. $\pi_u\left(\underline{\lambda}\right) \propto \left(\prod_{i=1}^{k} \lambda_i\right)^{-\frac{1}{2}}$

for the following Binomial distributions:

1. $n_1 = 10, p_1 = 0.4$ and $n_2 = 12, p_2 = 0.6$.

2. $n_1 = 20, p_1 = 0.4$ and $n_2 = 24, p_2 = 0.6$.

3. $n_1 = 40, p_1 = 0.4$ and $n_2 = 48, p_2 = 0.6$.

The priors denoted by π_s (1 and 3) are probability matching priors while those denoted by π_u (2 and 4) are Jeffreys' priors. The parameter of interest is $\psi = p_1 p_2$. The Poisson parameter $\lambda = np$.

Table 2. Frequentist Coverage Probabilities for 0.95 Posterior Quantile of $\psi = p_1 p_2$

	Binomial			**Poisson**		
$1-\alpha = 0.95$	**1**	**2**	**3**	**1**	**2**	**3**
n_1	10	20	40	10	20	40
p_1	0.4	0.4	0.4	0.4	0.4	0.4
λ_1				4	8	16
n_2	12	24	48	12	24	48
p_2	0.6	0.6	0.6	0.6	0.6	0.6
λ_2				7.2	14.4	28.8
λ_0				[4 7.2]	[8 14.4]	[16 28.8]
# x vectors	1 000	1 000	1 000	1 000	1 000	1 000
# $\lambda's$				1 000	1 000	1 000
π_s	0.953	0.954	0.95	0.933	0.933	0.949
π_u	0.926	0.944	0.946	0.913	0.917	0.949

$n = 1\,000$

From Tables 2 and 3 it is clear that the priors π_s are better than the Jeffreys' priors π_u in most of the situations. It is surprising that $\pi_s\left(\underline{\lambda}\right)$ is better than $\pi_u\left(\underline{p}\right)$ since $\pi_s\left(\underline{\lambda}\right)$ is the probability matching prior for the Poisson distribution. The latter will be a good approximation to the Binomial distribution if n is large and p is small. However the values used in Tables 2 and 3 are $p_1 = 0.4$ and $p_2 = 0.6$, which are quite large. As expected and although this is a limited experiment it seems that $\pi_s\left(\underline{p}\right)$ is the best prior of the four.

Table 3. Frequentist Coverage Probabilities for 0.05 Posterior Quantile of $\psi = p_1 p_2$

	Binomial			**Poisson**		
$\alpha = 0.05$	**1**	**2**	**3**	**1**	**2**	**3**
n_1	10	20	40	10	20	40
p_1	0.4	0.4	0.4	0.4	0.4	0.4
λ_1				4	8	16
n_2	12	24	48	12	24	48
p_2	0.6	0.6	0.6	0.6	0.6	0.6
λ_2				7.2	14.4	28.8
λ_0				[4 7.2]	[8 14.4]	[16 28.8]
# x vectors	1 000	1 000	1 000	1 000	1 000	1 000
# $\lambda' s$				1 000	1 000	1 000
π_s	0.048	0.047	0.052	0.048	0.055	0.064
π_u	0.027	0.031	0.042	0.038	0.043	0.054

$n = 1\,000$

4.3 A comparison of the Jeffreys', Uniform and Probability Matching priors for $\psi = p_1 p_2$

In this example a more extensive simulation study is done and coverage probabilities are obtained for $\psi = p_1 p_2$, the product of two Binomial parameters. For comparison purposes the following priors will be used:

1. The Jeffreys' prior: $\pi_u(p_1, p_2) = \pi_u(\underline{p}) \propto \prod_{i=1}^{2} p_i^{-\frac{1}{2}} (1 - p_i)^{-\frac{1}{2}}$.

2. The Uniform prior: $\pi(\underline{p}) \propto$ constant.

3. The Probability Matching prior: $\pi_s(\underline{p}) \propto \left\{ \sum_{i=1}^{2} (1 - p_i) p_i^{-1} \right\}^{\frac{1}{2}} \prod_{i=1}^{2} (1 - p_i)^{-1}$.

The parameter values for the Binomial distribution are $n_1 = n_2 = 10, n_1 = n_2 = 20$ and $p_i = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ (for $i = 1, 2$). The average length and standard deviation of the intervals are also given. The number of X variates are $1\,000$ and $n = 1\,000$.

In Table 4 summary statistics (averages over the nine possible values of the parameter p_2) are given for the coverage probabilities, mean lengths and standard deviations for the 90% credibility intervals of $\psi = p_1 p_2$.

From Table 4 it seems that the coverage probabilities for the Jeffreys' prior is in general somewhat smaller than 0.9 and that under coverage is larger for $n_1 = n_2 = 20$ than for $n_1 = n_2 = 10$. The uniform and probability matching priors on the other hand tend to give coverage probabilities larger than 0.9 and more so for the uniform prior. As can be expected the interval lengths and standard deviations are smaller for larger n.

Table 4. Coverage Rate of the 90% Credibility Intervals for $\psi = p_1 p_2$. Exact coverage probabilities (a), mean lengths (b), standard deviation (c). The values in this Table are averages over the nine possible values of p_2.

		$n_1 = n_2 = 10$			$n_1 = n_2 = 20$		
p_1		**Jeffreys'**	**Uniform**	**Probability matching**	**Jeffreys'**	**Uniform**	**Probability matching**
0.1	(a)	0.936	0.932	0.935	0.874	0.926	0.923
	(b)	0.156	0.175	0.174	0.111	0.119	0.120
	(c)	0.069	0.057	0.057	0.040	0.036	0.036
0.2	(a)	0.882	0.931	0.933	0.890	0.913	0.919
	(b)	0.210	0.217	0.218	0.155	0.157	0.158
	(c)	0.075	0.063	0.063	0.042	0.038	0.037
0.3	(a)	0.887	0.920	0.916	0.892	0.906	0.907
	(b)	0.251	0.252	0.252	0.188	0.187	0.187
	(c)	0.074	0.063	0.063	0.041	0.037	0.037
0.4	(a)	0.897	0.914	0.910	0.888	0.914	0.906
	(b)	0.288	0.282	0.284	0.214	0.212	0.211
	(c)	0.070	0.060	0.060	0.038	0.035	0.035
0.5	(a)	0.894	0.917	0.909	0.891	0.907	0.902
	(b)	0.315	0.309	0.307	0.236	0.232	0.232
	(c)	0.065	0.056	0.057	0.036	0.033	0.033
0.6	(a)	0.903	0.913	0.910	0.887	0.906	0.901
	(b)	0.338	0.329	0.329	0.252	0.249	0.249
	(c)	0.060	0.051	0.052	0.033	0.030	0.030
0.7	(a)	0.893	0.909	0.905	0.890	0.906	0.902
	(b)	0.357	0.348	0.347	0.267	0.264	0.263
	(c)	0.056	0.047	0.048	0.031	0.028	0.028
0.8	(a)	0.897	0.902	0.898	0.887	0.900	0.902
	(b)	0.372	0.363	0.363	0.278	0.275	0.275
	(c)	0.056	0.044	0.045	0.030	0.027	0.027
0.9	(a)	0.899	0.896	0.900	0.888	0.901	0.898
	(b)	0.381	0.375	0.376	0.286	0.284	0.284
	(c)	0.057	0.043	0.042	0.033	0.027	0.027
Overall Mean	(a)	0.899	0.915	0.913	0.887	0.909	0.907
	(b)	0.296	0.295	0.294	0.221	0.220	0.220
	(c)	0.065	0.054	0.054	0.036	0.032	0.032

It also seems that the probability matching prior gives the best results for $0.3 \leq p_i \leq 0.7, (i = 1, 2)$. This also explains the good performance of the probability matching priors in Tables 2 and 3. In Table 5 the overall averages are given for $n_1 = n_2 = 10$ and $n_1 = n_2 = 20$ for $p_i = 0.3, 0.4, 0.5, 0.6$ and $0.7, (i = 1, 2)$.

From Table 5 it can be seen that the probability matching prior is somewhat better than the uniform and Jeffreys' priors. We will conclude by saying that all three priors are doing well for attaining the nominal coverage probabilities. In general the differences between the priors are quite small.

Table 5. Average Coverage probabilities of the 90% Credibility Intervals for $n_1 = n_2 = 10$ and $n_1 = n_2 = 20$ for $p_i = 0.3, 0.4, 0.5, 0.6$ and $0.7, (i = 1, 2)$.

$n_1 = n_2 = 10$			$n_1 = n_2 = 20$		
Jeffreys'	**Uniform**	**Probability matching**	**Jeffreys'**	**Uniform**	**Probability matching**
0.890	0.910	0.905	0.891	0.907	0.900

5. Bayesian Confidence Interval Estimation for a Linear Function of Binomial Proportions

Due to its important practical value, confidence interval construction for a linear function of Binomial proportions has received some attention recently (Price & Bonett, 2004; Tebbs & Roths, 2008). In the first part of this section the probability matching prior for a linear function of Binomial proportions will therefore be derived and in the latter part Bayesian confidence intervals will be constructed for the difference between two Binomial proportions.

Estimating the difference between two binomial proportions is a problem that occurs regularly in practice. There are some asymptotic procedures available for the construction of confidence intervals for the difference. A number of authors have studied the performance of these asymptotic procedures in circumstances where the samples are small. Some of them are Agresti & Caffo (2000), Beal (1987), Newcombe (1998) and Zhou *et al.* (2004). According to Roths & Tebbs (2006) asymptotic intervals are generally preferred to exact intervals. The reason for this, is that asymptotic intervals are often much easier to calculate than exact intervals and they can also produce acceptable results without wasteful conservatism. Roths & Tebbs (2006) showed how their intervals can be used adaptively in experiments conducted in stages over time, they concentrated on samples that are small.

5.1 The Probability Matching Prior for a Linear Combination of Binomial Proportions

The procedure of Datta & Ghosh (1995) will be used to derive the probability matching prior. The following theorem can be proved.

Theorem 5.1 The probability matching prior for $\theta = \sum_{i=1}^{k} \alpha_i p_i$ a linear combination of Binomial proportions is given by

$$\widetilde{\pi}_s(p_1, p_2, \ldots, p_k) \propto \left\{ \sum_{i=1}^{k} \alpha_i^2 p_i (1 - p_i) \right\}^{\frac{1}{2}} \prod_{i=1}^{2} p_i^{-1} (1 - p_i)^{-1}.$$

Proof. As in Theorem 2.1, the inverse of the Fisher Information matrix is given by

$$F^{-1}(\underline{p}) = diag\left[\; p_1(1-p_1) \quad p_2(1-p_2) \quad \ldots \quad p_k(1-p_k) \;\right].$$

We are interested in a probability matching prior for $t\left(\underline{p}\right)=\theta=\sum_{i=1}^{k}\alpha_i p_i$, a linear combination of k Binomial parameters.

Now

$$\begin{aligned}\nabla_t'\left(\underline{p}\right) &= \begin{bmatrix} \frac{\partial t(\underline{p})}{\partial p_1} & \frac{\partial t(\underline{p})}{\partial p_2} & \cdots & \frac{\partial t(\underline{p})}{\partial p_k}\end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{bmatrix}.\end{aligned}$$

Also

$\nabla_t'\left(\underline{p}\right)F^{-1}\left(\underline{p}\right)=\begin{bmatrix} \alpha_1 p_1\left(1-p_1\right) & \alpha_2 p_2\left(1-p_2\right) & \cdots & \alpha_k p_k\left(1-p_k\right)\end{bmatrix}$

and

$\nabla_t'\left(\underline{p}\right)F^{-1}\left(\underline{p}\right)\nabla_t\left(\underline{p}\right)=\sum_{i=1}^{k}\alpha_i^2 p_i\left(1-p_i\right)$.

Further

$$\begin{aligned}\eta'\left(\underline{p}\right) &= \frac{\nabla_t'\left(\underline{p}\right)F^{-1}\left(\underline{p}\right)}{\sqrt{\nabla_t'\left(\underline{p}\right)F^{-1}\left(\underline{p}\right)\nabla_t\left(\underline{p}\right)}} \\ &= \begin{bmatrix} \eta_1\left(\underline{p}\right) & \eta_2\left(\underline{p}\right) & \cdots & \eta_k\left(\underline{p}\right)\end{bmatrix}\end{aligned}$$

where $\eta_i\left(\underline{p}\right)=\frac{\alpha_i p_i(1-p_i)}{\sqrt{\sum_{i=1}^{k}\alpha_i^2 p_i(1-p_i)}}$ $\quad\quad (i=1,\ldots,k)$.

The prior $\pi\left(\underline{p}\right)$ is a probability matching prior if and only if the differential equation $\sum_{i=1}^{k}\frac{\partial}{\partial p_i}\left\{\eta_i\left(\underline{p}\right)\pi\left(\underline{p}\right)\right\}=0$ is satisfied.

The differential equation will be satisfied if $\pi\left(\underline{p}\right)$ is

$$\widetilde{\pi}_s\left(\underline{p}\right)\propto\left\{\sum_{i=1}^{k}\alpha_i^2 p_i\left(1-p_i\right)\right\}^{\frac{1}{2}}\prod_{i=1}^{k}p_i^{-1}\left(1-p_i\right)^{-1}. \blacksquare$$

If $k=2, \alpha_1=1, \alpha_2=-1$ and $\alpha_3=\alpha_4=\ldots=\alpha_k=0$, then $\theta=p_1-p_2$ and the posterior distribution in the case of the probability matching prior is

$$\widetilde{\pi}_s\left(p_1,p_2\,|data\right) \propto\left\{\sum_{i=1}^{2}p_i\left(1-p_i\right)\right\}^{\frac{1}{2}}\prod_{i=1}^{2}p_i^{x_i-1}\left(1-p_i\right)^{n_i-x_i-1}. \tag{3}$$

for $0\leq p_i\leq 1$.

If $k=1$ and $\alpha=1$, $\widetilde{\pi}_s\left(p\right)$ becomes the Jeffreys' prior.

Theorem 5.2 The posterior distribution defined in equation 3 is a proper distribution if $0<x_i<n_i$.

Proof. Since

$$\left\{\sum_{i=1}^{2} p_i\left(1-p_i\right)\right\}^{\frac{1}{2}} < \sum_{i=1}^{2} p_i^{\frac{1}{2}}\left(1-p_i\right)^{\frac{1}{2}}$$

it follows that

$$\begin{aligned}\widetilde{\pi}_s\left(p_1,p_2\,|data\right) &\propto \left\{\sum_{i=1}^{2} p_i\left(1-p_i\right)\right\}^{\frac{1}{2}} \prod_{i=1}^{2} p_i^{x_i-1}\left(1-p_i\right)^{n_i-x_i-1} \\ &< \left\{\sum_{i=1}^{2} p_i^{\frac{1}{2}}\left(1-p_i\right)^{\frac{1}{2}}\right\} \prod_{i=1}^{2} p_i^{x_i-1}\left(1-p_i\right)^{n_i-x_i-1}.\end{aligned}$$

Now
$\int_0^1\int_0^1\left\{\sum_{i=1}^{2} p_i^{\frac{1}{2}}\left(1-p_i\right)^{\frac{1}{2}}\right\}\left\{\prod_{i=1}^{2} p_i^{x_i-1}\left(1-p_i\right)^{n_i-x_i-1}\right\}dp_1dp_2$ will converge if $x_i>0$ and $x_i<n_i$,$(i=1,2)$.■

5.2 Simulation Study

In this section an extensive simulation study will be done and coverage probabilities will be obtained for $\theta = p_1 - p_2$. The two Bayesian methods (Jeffreys' and probability matching priors) will be compared with known classical procedures. The Jeffreys' prior is defined as

$\widetilde{\pi}_u\left(\underline{p}\right) = \widetilde{\pi}_u\left(p_1,p_2\right) \propto \prod_{i=1}^{2} p_i^{-\frac{1}{2}}\left(1-p_i\right)^{-\frac{1}{2}}$.

When using the Jeffreys' prior the joint posterior distribution is the product of independent $Beta\left(x_i+\frac{1}{2}, n_i+x_i+\frac{1}{2}\right)$, $(i=1,2)$, variates which is denoted by

$$\widetilde{\pi}_u\left(p_1,p_2\,|data\right) \propto \left\{\prod_{i=1}^{2} Beta\left(x_i+\frac{1}{2}, n_i+x_i+\frac{1}{2}\right)\right\}^{-1} \prod_{i=1}^{2} p_i^{x_i-\frac{1}{2}}\left(1-p_i\right)^{n_i-x_i-\frac{1}{2}}.$$

The credibility (Bayesian confidence) intervals for the probability matching are obtained by using an adapted weighted Monte Carlo method as described in Section . The number of X vectors is again 1 000 and $n = 1000$.

Tables 6 and 7 contain coverage probabilities, mean lengths and conditional mean length ratios for the six intervals discussed by Roths & Tebbs (2006) as well as the full (real) Bayesian procedures, by using Jeffreys' prior (Bayes (J)) and the probability matching prior (Bayes (PMP)), when $n_1 = n_2 = 10$ and $n_1 = n_2 = 20$, respectively, for a number of choices for p_1 and p_2. The nominal confidence level is $1-\alpha = 0.95$. The conditional mean length ratio is the ratio of the mean lengths for cases where the differences are covered and when they are not covered. A small value of the conditional mean length ratio is desirable.

Table 6. Exact coverage probabilities (a), mean lengths (b), and conditional mean length ratios (c) for $n_1 = n_2 = 10$. The nominal level is 0.95.WAL, Wald; AGC, Agresti-Caffo; HAL, Haldane; JFP, Jeffreys-Perks; MLE, Beal-MLE; MOM, Beal-MOM; Bayes (J), Bayesian procedure using Jeffreys' prior; Bayes (PMP), Bayesian procedure using the probability matching prior.

p_1	p_2		**WAL**	**AGC**	**HAL**	**JFP**	**MLE**	**MOM**	**Bayes (J)**	**Bayes (PMP)**
0.1	0.1	(a)	0.95000	0.99100	0.99100	0.99100	0.99100	0.99100	0.97000	0.99100
		(b)	0.45600	0.57800	0.41900	0.52300	0.47700	0.41900	0.53600	0.56840
		(c)	0.77500	0.90000	0.72700	0.86100	0.78200	0.72700	0.93200	0.94473
0.1	0.3	(a)	0.93900	0.96800	0.94300	0.94300	0.94600	0.94300	0.95000	0.97300
		(b)	0.63200	0.65700	0.58500	0.62600	0.63600	0.58500	0.63200	0.63854
		(c)	1.34800	1.10000	1.35800	1.18300	1.42000	1.35700	1.11900	1.11850
0.1	0.5	(a)	0.91100	0.96300	0.91500	0.93000	0.94900	0.91500	0.93900	0.95400
		(b)	0.67800	0.68200	0.64000	0.65700	0.66800	0.64000	0.65300	0.65098
		(c)	1.27300	1.02800	1.16900	1.09100	1.10400	1.16500	1.11000	1.01310
0.1	0.7	(a)	0.91500	0.94500	0.94500	0.94500	0.96000	0.94500	0.95600	0.91700
		(b)	0.61900	0.65600	0.61800	0.62400	0.62700	0.61900	0.61400	0.61652
		(c)	1.74100	0.99000	0.99000	0.99000	0.97600	0.98100	0.95100	0.87559
0.1	0.9	(a)	0.87000	0.95700	0.94900	0.95700	0.95700	0.94900	0.94600	0.87000
		(b)	0.41500	0.56800	0.51600	0.51800	0.51800	0.51700	0.48900	0.52713
		(c)	9.76600	0.77500	0.73700	0.71900	0.71300	0.73800	0.70900	0.80078
0.3	0.3	(a)	0.90500	0.96300	0.96300	0.96300	0.96300	0.96300	0.90900	0.95400
		(b)	0.75600	0.72700	0.69600	0.71400	0.73900	0.69600	0.71900	0.69898
		(c)	1.11500	1.08900	1.11800	1.11600	1.15800	1.10800	1.13400	1.11780
0.3	0.5	(a)	0.92200	0.96400	0.94900	0.94900	0.95800	0.94900	0.95400	0.95700
		(b)	0.79400	0.75000	0.73400	0.74100	0.75700	0.73500	0.73900	0.71538
		(c)	1.16200	1.09000	1.11700	1.11100	1.14700	1.11600	1.13600	1.07780
0.3	0.7	(a)	0.93200	0.95500	0.94100	0.94100	0.95700	0.94100	0.93500	0.95000
		(b)	0.75400	0.72700	0.71000	0.71300	0.72000	0.71000	0.70700	0.68957
		(c)	1.37700	0.99500	1.06300	1.06000	1.11900	1.06100	1.09700	0.96325
0.5	0.5	(a)	0.91200	0.95800	0.95800	0.95800	0.95800	0.95800	0.94400	0.96100
		(b)	0.83000	0.77100	0.76400	0.76800	0.78000	0.76500	0.76300	0.73764
		(c)	1.12900	1.10400	1.13300	1.13200	1.14600	1.13400	1.11700	1.11680
Overall Mean		(a)	0.917	0.963	0.950	0.953	0.960	0.950	0.945	0.948
		(b)	0.659	0.680	0.631	0.654	0.658	0.632	0.650	0.649
		(c)	2.187	1.008	1.046	1.029	1.063	1.043	1.034	1.003

The differences among all the intervals are not too large, the only exception to this is the Wald interval. The results from the Bayesian procedures compare well with the other methods.

Table 7. Exact coverage probabilities (a), mean lengths (b), and conditional mean length ratios (c) for $n_1 = n_2 = 20$. The nominal level is 0.95.WAL, Wald; AGC, Agresti-Caffo; HAL, Haldane; JFP, Jeffreys-Perks; MLE, Beal-MLE; MOM, Beal-MOM; Bayes (J), Bayesian procedure using Jeffreys' prior; Bayes (PMP), Bayesian procedure using the probability matching prior.

p_1	p_2		**WAL**	**AGC**	**HAL**	**JFP**	**MLE**	**MOM**	**Bayes (J)**	**Bayes (PMP)**
0.1	0.1	(a)	0.96000	0.98800	0.96100	0.98300	0.96100	0.96100	0.94800	0.97400
		(b)	0.35200	0.39600	0.33600	0.37000	0.36700	0.33600	0.38000	0.39128
		(c)	0.92200	0.93300	0.92500	0.89000	0.96900	0.92500	1.01300	0.96505
0.1	0.3	(a)	0.93500	0.96000	0.95200	0.95200	0.95200	0.95200	0.94100	0.96000
		(b)	0.46400	0.47300	0.44600	0.46000	0.46900	0.44600	0.46200	0.46414
		(c)	1.08400	1.05400	1.10400	1.08600	1.07200	1.10400	1.05700	1.07410
0.1	0.5	(a)	0.93700	0.95400	0.94300	0.95000	0.95000	0.94300	0.95200	0.95000
		(b)	0.49600	0.49600	0.48000	0.48600	0.49200	0.48000	0.48200	0.48345
		(c)	1.09900	1.03800	1.04200	1.05300	1.04200	1.03800	1.06500	0.98793
0.1	0.7	(a)	0.91300	0.95500	0.93300	0.93300	0.93700	0.93300	0.94500	0.94200
		(b)	0.46400	0.47300	0.45700	0.45900	0.46100	0.45900	0.45700	0.45785
		(c)	1.25200	0.93200	1.02000	1.01900	1.00300	1.02300	1.02900	0.90632
0.1	0.9	(a)	0.91300	0.95800	0.94300	0.94300	0.94300	0.94300	0.94100	0.90900
		(b)	0.34200	0.39400	0.36800	0.36900	0.36900	0.36900	0.36100	0.37578
		(c)	2.41900	0.78100	0.89900	0.89900	0.89800	0.89900	0.87900	0.80140
0.3	0.3	(a)	0.93100	0.95000	0.94700	0.94700	0.95500	0.94700	0.93800	0.94800
		(b)	0.55200	0.53800	0.52800	0.53400	0.55000	0.52800	0.53400	0.52510
		(c)	1.05900	1.04000	1.05500	1.05300	1.08300	1.05500	1.06600	1.04840
0.3	0.5	(a)	0.94200	0.95200	0.94600	0.94600	0.95100	0.94600	0.94900	0.95800
		(b)	0.57900	0.55900	0.55500	0.55600	0.56500	0.55500	0.55600	0.54606
		(c)	1.07900	1.04300	1.05500	1.05400	1.06600	1.05500	1.06300	1.03000
0.3	0.7	(a)	0.92800	0.94400	0.94400	0.94400	0.94400	0.94400	0.95000	0.94800
		(b)	0.55200	0.53800	0.53300	0.53300	0.53600	0.53300	0.52900	0.52428
		(c)	1.15400	1.03300	1.03900	1.03900	1.03700	1.03900	1.01900	1.03680
0.5	0.5	(a)	0.91900	0.95700	0.95700	0.95700	0.96100	0.95700	0.93800	0.95600
		(b)	0.60400	0.57800	0.57800	0.57800	0.58300	0.57800	0.57700	0.56406
		(c)	1.05600	1.06000	1.06800	1.06800	1.07200	1.06800	1.07300	1.06860
Overall Mean		(a)	0.931	0.958	0.947	0.951	0.950	0.947	0.945	0.949
		(b)	0.489	0.494	0.476	0.483	0.488	0.476	0.482	0.481
		(c)	1.236	0.990	1.023	1.018	1.027	1.023	1.029	0.991

5.3 Example

In this section the adaptability of the intervals are shown in situations where data are collected in multiple stages. To illustrate this consider the data from Ornaghi *et al.* (1999) given in Table 8. The stages correspond to different dates on which insects were collected during maize planting season in Argentina (from October to November). The goal of this experiment was to assess if male and female insects transmit the Mal de Rio Cuarto virus to susceptible maize plants at similar rates.

Table 8. Number of test plants and numbers of virus-infected plants collected on five different dates during the maize plant season.

Stage	Gender	Number of test plants	Number infected
1	M	29	9
	F	31	5
2	M	57	4
	F	57	7
3	M	57	8
	F	57	16
4	M	24	2
	F	24	3
5	M	24	3
	F	24	2

Assume that, at a specific stage, the researchers want to estimate the difference $p_1 - p_2$, where p_1 is equal to the proportion infected plants for male insects and p_2 is the proportion infected plants for female insects. In Table 9 the 95% confidence intervals for $\theta = p_1 - p_2$ are given for the six methods described by Roths & Tebbs (2006) and in Table 10 the 95% Bayesian confidence intervals are given for the Jeffreys' prior, uniform prior and probability matching prior.

Table 9. 95% Confidence intervals for the difference in disease transmission probabilities among male and female insects.

Stage	Interval	Lower limit	Upper limit	Length
1	WAL	-0.063	0.361	0.425
	AGC	-0.070	0.351	0.421
	HAL	-0.065	0.347	0.412
	JFP	-0.067	0.350	0.417
	MLE	-0.072	0.354	0.427
	MOM	-0.065	0.347	0.412
2	WAL	-0.161	0.055	0.216
	AGC	-0.163	0.062	0.225
	HAL	-0.157	0.055	0.212
	JFP	-0.160	0.059	0.219
	MLE	-0.163	0.061	0.224
	MOM	-0.157	0.055	0.212
3	WAL	-0.288	0.007	0.295
	AGC	-0.283	0.012	0.295
	HAL	-0.281	0.009	0.290
	JFP	-0.282	0.011	0.293
	MLE	-0.283	0.012	0.295
	MOM	-0.281	0.009	0.290

According to Roths & Tebbs (2006) if suitable computing facilities are available, they would recommend either the Jeffreys-Perks or Beal-MLE interval, because their coverage probabilities are closer to the nominal level than those for the Haldane and Beal-MOM intervals and are not as conservative as the

Agresti-Caffo interval.

Table 10. 95% Bayesian Confidence intervals for the difference in disease transmission probabilities among male and female insects. Using the Jeffreys' prior, uniform prior and probability matching prior.

Stage	Interval	Lower limit	Upper limit	Length
1	Jeffreys'	-0.073	0.366	0.439
	Uniform	-0.057	0.348	0.405
	PMP	-0.078	0.346	0.423
2	Jeffreys'	-0.175	0.056	0.231
	Uniform	-0.177	0.071	0.248
	PMP	-0.163	0.071	0.234
3	Jeffreys'	-0.293	0.004	0.297
	Uniform	-0.267	0.016	0.284
	PMP	-0.280	0.011	0.290
4	Jeffreys'	-0.227	0.138	0.365
	Uniform	-0.225	0.156	0.381
	PMP	-0.219	0.144	0.363
5	Jeffreys'	-0.138	0.223	0.361
	Uniform	-0.136	0.232	0.369
	PMP	-0.141	0.228	0.368

From Table 10 it can be seen that the Bayesian confidence intervals when using the Jeffreys' prior compares well with the other methods. The lower limits of the intervals are however in general somewhat smaller and the interval lengths somewhat larger than those of the methods suggested by Roths & Tebbs (2006). From our simulation studies (Tables 6 and 7) it was clear that there is not much to choose between the Jeffreys' and probability matching priors.

6. Conclusion

In this paper probability matching priors for the product of k independent Binomial rates, i.e. $\psi = \prod_{i=1}^{k} p_i^{\alpha_i}$ and also for a linear combination of Binomial rates, $\theta = \sum_{i=1}^{k} \alpha_i p_i$, were derived. Limited simulation studies have shown that the probability matching prior achieves its sample frequentist coverage results somewhat better than in the case of the Jeffreys' prior.

References

AGRESTI, A., & CAFFO, B. 2000. Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures. *The American Statistician*, **54**(4), 280–288.

BAYARRI, M. J., & BERGER, J. O. 2004. The Interplay of Bayesian and Frequentist Analysis. *Statistical Science*, **19**(1), 58–80.

BEAL, S. L. 1987. Asymptotic Confidence Intervals for the Difference between Two Binomial Parameters for Use with Small Samples. *Biometrics*, **43**(4), 941–950.

CHEN, M-H., & SHAO, Q-M. 1999. Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *Journal of Computational and Graphical Statistics*, **8**(1), 69–92.

DATTA, G.S., & GHOSH, J.K. 1995. On Priors Providing Frequentist Validity for Bayesian Inference. *Biometrika*, **82**(1), 37–45.

GUTTMAN, I., & MENZEFRIEKE, U. 2003. Posterior Distributions for Functions of Variance Components. *Sociedad de Estadistica e Investigacio Operativa Test*, **12**, 115 – 123.

HARRIS, B. 1971. Hypothesis Testing and Confidence Intervals for Products and Quotients of Poisson Parameters with Applications to Reliability. *Journal of the American Statistical Association*, **66**(335), 609–613.

KIM, H-J. 2006. On Bayesian Estimation of the Product of Poisson Rates with Application to Reliability. *Communications in Statistics - Simulation and Computation*, **35**, 47 – 59.

LI, K-H. 2007. Pool size selection for the sampling/importance resampling algorithm. *Statistica Sinica*, **17**(3), 895 – 907.

MADANSKY, A. 1965. Approximate Confidence Limits for the Reliability of Series and Parallel Systems. *Technometrics*, **7**(4), 495–503.

NEWCOMBE, R.G. 1998. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Statistics in Medicine*, **17**(8), 873–890.

ORNAGHI, J.A., MARCH, G.J., BOITO, G.T., MARINELLI, A., BEVIACQUA, J.E., GIUGGIA, J., & LENARDON, S.L. 1999. Infectivity in natural populations of Delphacodes kuscheli vector of "Mal Rio Cuarto" virus. *Maydica*, **44**, 219 – 223.

PRICE, R.M., & BONETT, D.G. 2004. An improved confidence interval for a linear function of binomial proportions. *Computational Statistics & Data Analysis*, **45**, 449 – 456.

ROTHS, S.A., & TEBBS, J.M. 2006. Revisiting Beal's Confidence Intervals for the Difference of Two Binomial Proportions. *Communications in Statistics - Theory and Methods*, **35**(9), 1593 – 1609.

RUBIN, D.B. 1984. Bayesian Justifiable and Relevant Frequency Calculation for the Applied Statistician. *The Annals of Statistics*, **12**, 1151 – 1172.

SEVERINI, T.A., MUKERJEE, R., & GHOSH, M. 2002. On An Exact Probability Matching Property of Right-invariant Priors. *Biometrika*, **89**, 952 – 957.

SKARE, O., BOLVIKEN, E., & HOLDEN, L. 2003. Improved Sampling-Importance Resampling and Reduced Bias Importance Sampling. *Scandinavian Journal of Statistics*, **30**(4), 719–737.

SMITH, A. F. M., & GELFAND, A. E. 1992. Bayesian Statistics without Tears: A Sampling-Resampling Perspective. *The American Statistician*, **46**(2), 84–88.

TEBBS, J.M., & ROTHS, S.A. 2008. New large-sample confidence intervals for a linear combination of binomial proportions. *Journal of Statistical Planning and Inference*, **138**, 1884 – 1893.

WOLPERT, R.L. 2004. A Conversation with James O Berger. *Statistical Science*, **91**, 205 – 218.

ZHOU, X-H., TSAO, M., & QIN, Q. 2004. New intervals for the difference between two independent binomial proportions. *Journal of Statistical Planning and Inference*, **123**(1), 97 – 115.