Scientific Research Publishing

Open Journal of Statistics, 2017, 7, 1-15
http://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Inferences on the Difference of Two Proportions: A Bayesian Approach

Thu Pham-Gia[1], Nguyen Van Thin[2], Phan Phuc Doan[2]

[1]Department of Mathematics and Statistics, Université de Moncton, New Brunswick, Canada
[2]Faculty of Mathematics and Computer Science, Hochiminh University of Science, Ho Chi Minh, Vietnam
Email: thu.pham-gia@umoncton.ca, nvthin@hcmus.edu.vn, ppdoan@hcmus.edu.vn

How to cite this paper: Pham-Gia, T., Thin, N.V. and Doan, P.P. (2017) Inferences on the Difference of Two Proportions: A Bayesian Approach. *Open Journal of Statistics*, **7**, 1-15.
https://doi.org/10.4236/ojs.2017.71001

Received: July 13, 2016
Accepted: February 6, 2017
Published: February 9, 2017

Abstract

Let $\pi = \pi_1 - \pi_2$ be the difference of two independent proportions related to two populations. We study the test $H_0 : \pi \geq 0$ against different alternatives, in the Bayesian context. The various Bayesian approaches use standard beta distributions, and are simple to derive and compute. But the more general test $H_0 : \pi \geq \eta$, with $\eta > 0$, requires more advanced mathematical tools to carry out the computations. These tools, which include the density of the difference of two general beta variables, are presented in the article, with numerical examples for illustrations to facilitate comprehension of results.

Keywords

Proportion, Convolution, Normal, Beta, Bayesian, Critical Value, Appell's Function

1. Introduction

For two independent proportions π_1 and π_2, their difference is frequently encountered in the frequentist statistical literature, where tests, or confidence intervals, for $\pi_1 - \pi_2$ are well accepted notions in theory and in practice, although most frequently, the case under study is the equality, or inequality of these proportions. For the Bayesian approach, Pham-Gia and Turkkan ([1] and [2]) have considered the case of independent, and dependent proportions for inferences, and also in the context of sample size determination [3].

But testing $\pi_1 = \pi_2$ is only a special case of testing $H_0 : \pi_1 - \pi_2 \leq \eta$, with η being a positive constant value, which is much less frequently dealt with. In Section 2 we recall the unconditional approaches to testing H_0 based on the maximum likelihood estimators of the two proportions and normal approximations. A new exact approach not using normal approximation has been developed by our group and will be presented elsewhere. Fisher's exact test is also re-

called here, for comparison purpose. The Bayesian approach to testing the equality of two proportions and the computation of credible intervals are given in Section 3. The Bayesian approach using the general beta distributions is given in Section 4. All related problems are completely solved, thanks to some closed form formulas that we have established in earlier papers.

2. Testing the Equality of Two Proportions

2.1. Test Using Normal Approximation

As stated before, taking $\eta = 0$ we have a test for equality between two proportions. Several well-known methods are presented in the literature. For example, the conditional test is usually called Fisher's exact test, and is based on the hypergeometric distribution. It is used when the sample size is small. Pearson's Chi-square test using Yates correction is usually used for intermediary sample size while Pearson's Chi-square is used for large samples. Their appropriateness is discussed in D'Agostino *et al.* [4]. Normal approximation methods are based on formulas using estimated values of the mean and the variance of the two populations. For example, we have

$T_1 = \dfrac{X_1/n_1 - X_2/n_2}{\left[\left(X_1/n_1\right)\left(1-X_1/n_1\right)/n_1 + \left(X_2/n_2\right)\left(1-X_2/n_2\right)/n_2\right]^{1/2}}$, and the pooled version

$T_2 = \dfrac{X_1/n_1 - X_2/n_2}{\left[\left(X_1+X_2\right)/\left(n_1+n_2\right)\left(\left(1-\left(X_1+X_2\right)\right)/\left(n_1+n_2\right)\right)\left(1/n_1+1/n_2\right)\right]^{1/2}}$, both being approximately $N(0,1)$ under $H_0 : \pi_1 \le \pi_2$. Cressie [5] gives conditions under which T_2 is better than T_1, in terms of power. Previously, Eberhardt and Fligner [6] studied the same problem for a bilateral test.

Numerical Example 1

To investigate its proportions of customers in two separate geographic areas of the country, a company picks a random sample of 25 shoppers in area *A*, in which 17 are found to be its customers. A similar random sample of 20 shoppers in area *B* gives 8 customers. We wish to test the hypothesis that $H_0 : \pi_1 \le \pi_2$ against $H_1 : \pi_1 > \pi_2$.

We have here the observed value of $T_1 = 1.9459$ and of $T_2 = 1.8783$ which lead, in both cases, to the rejection of H_0 at significance level 5% (the critical value is 1.64) for $H_1 : \pi_1 > \pi_2$.

2.2. Fisher's Exact Test

Under H_0 the number of successes coming from population 1 has the $\mathrm{Hyp}\left(n_1+n_2, t = x_1+x_2, n_1, x\right)$ distribution. The argument is that, in the combined sample of size n_1+n_2, with x_1 successes from population 1 out of the total number of successes $t = x_1 + x_2$, the number of *x* successes coming from population 1 is a hypergeometric variable.

To compute the significance of the observation we have to compute several tables corresponding to more extreme results than the observed table. It is known that the conditional test is less powerful than the unconditional one.

Numerical Example 2

We use the same data as in numerical example 1 to test $H_0 : \pi_A = \pi_B$ vs $H_1 : \pi_A > \pi_B$ *i.e.* the proportion of customers in area *A* is significantly higher than the one in area *B*. We have **Table 1**:

the observed data $(x_B = 8)$, and also cases more extreme, which means $x_B = 0, 1, 2, \cdots, 7$. The *p*-value of the test is hence

$$p\text{-value} = \sum_{x_B=0}^{8} \frac{\binom{25}{25-x_B}\binom{20}{x_B}}{\binom{45}{25}} = 0.0542 .$$

Although technically not significant at the 5% level, this result shows that the proportion of customers in area *B* can practically be considered as lower than the one in area *A*, in agreement with the frequentist test.

REMARK: The problem is often associated with a 2 × 2 table where there are three possibilities: constant column sums and row sums, one set constant the other variable and both variables. Other measures can then be introduced (e.g. Santner and Snell [7]). A Bayesian approach has been carried out by several authors, e.g. Howard [8] and also Pham Gia and Turkkan [2], who computed the credible intervals for several of these measures.

3. The Bayesian Approach

In the estimation of the difference of two proportions the Bayesian approach certainly plays an important role. Agresti and Coull [9] provide some interesting remarks on various approaches.

Again, let $\pi = \pi_1 - \pi_2$. Using the Bayesian approach will certainly encounter some serious computational difficulties if we do not have a closed form expression for the density of the difference of two independently beta distributed random variables. Such an expression has been obtained by the first author some time ago and is recalled below.

3.1. Bayesian Test on the Equality of Two Proportions

Let us recall first the following theorem:

Theorem 1: Let $\pi_i \sim \text{beta}(\alpha_i, \beta_i)$, for $i = 1, 2$ be two independent beta distributed random variables with parameters (α_1, β_1) and (α_2, β_2), respectively. Then the difference $\pi = \pi_1 - \pi_2$ has density defined on $(-1, 1)$ as follows:

$$p_\pi(x) = \begin{cases} B(\alpha_2, \beta_1)\, x^{\beta_1+\beta_2-1}(1-x)^{\alpha_2+\beta_1-1} \\ \quad F_1\left(\beta_1, \alpha_1+\alpha_2+\beta_1+\beta_2-2, 1-\alpha_1; \beta_1+\alpha_2; \left((1-x), 1-x^2\right)\right) \Big/ A, \; 0 \le x < 1 \\ B(\alpha_1+\alpha_2-1; \beta_1+\beta_2-1) \Big/ A, \; x = 0, \text{ if } \alpha_1+\alpha_2 > 1, \; \beta_1+\beta_2 > 1 \\ B(\alpha_1, \beta_2)(-x)^{\beta_1+\beta_2-1}(1+x)^{\alpha_1+\beta_2-1} \\ \quad F_1\left(\beta_2, 1-\alpha_2, 1-\alpha_2; \alpha_1+\alpha_2+\beta_1+\beta_2-2, \alpha_1+\beta_2; 1-x^2, 1+x\right) \Big/ A, \; -1 \le x < 0 \end{cases} \quad (1)$$

$$A = B(\alpha_1, \beta_1)\, B(\alpha_2, \beta_2)$$

$F_1(.)$ is Appell's first hypergeometric function, which is defined as

$$F_1(a,b_1,b_2;c;x_1,x_2)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{a^{[i+j]}}{c^{[i+j]}}\frac{b_1^{[i]}b_2^{[j]}x_1^i x_2^j}{i!\,j!} \tag{2}$$

where $a^{[b]}=a(a+1)\cdots(a+b-1)$. This infinite series is convergent for $|x_1|<1$ and $|x_2|<1$, where, as shown by Euler, it can also be expressed as a convergent integral:

$$\frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\int_0^1 u^{a-1}(1-u)^{c-a-1}(1-ux_1)^{-b_1}(1-ux_2)^{-b_2}\,\mathrm{d}u \tag{3}$$

which converges for $c-a>0$, $a>0$. In fact, Pham-Gia and Turkkan [1] established the expression of the density of the difference using (3) directly and not the series. Hence, the infinite series (5) can be extended outside the two circles of convergence, by analytic continuation, where it is also denoted by $F_1(.)$.

Here, we denote the above density (1) by $\pi\sim\psi(\alpha_1,\beta_1,\alpha_2,\beta_2)$.

Proof: See Pham-Gia and Turkkan [1].

The prior distribution of π is hence $\psi(\alpha_1,\beta_1,\alpha_2,\beta_2)$, obtained from the two beta priors. Various approaches in Bayesian testing are given below.

Bayesian Testing Using a Significance Level

While frequentist statistics frequently does not test $H_0:\pi\le\eta$ vs. $H_1:\pi>\eta$, for $\eta>0$ and limits itself to the case $\eta=0$, Bayesian statistics can easily do it.

a) One-sided test:

Proposition 1: To perform the above test at the 0.05 significance level, using the two independent samples $\{X_{1,i}\}_{i=1}^{n_1}$ and $\{X_{2,i}\}_{i=1}^{n_2}$, we compute $p_{\pi_1-\pi_2}(\pi_1-\pi_2|\alpha_1^*,\beta_1^*,\alpha_2^*\beta_2^*)$, where $\alpha_i^*=\alpha_i+x_i$ and $\beta_i^*=\beta_i+n_i-x_i$, $i=1,2$. This expression of the posterior density of π, obtained by the conjugacy of binomial sampling with the beta prior, will allow us to compute $P(\pi>\eta)$ and compare it with the significance level α.

For example, as in the frequentist example of Section 2.1, we consider $n_1=25$, $x_1=17$, $n_2=20$, $x_2=8$ and use two non-informative beta priors, that is, $\mathrm{Beta}(0.5,0.5)$.

We note first that $\hat{\pi}_1=17/25=0.68$, $\hat{\pi}_2=8/20=0.40$, giving $\hat{\pi}=0.28$.

We obtain the prior and posterior distributions of π_1 and π_2 (**Figure 1**). We wish to test:

$$H_0:\pi\le 0.35 \quad \text{vs} \quad H_1:\pi>0.35 \tag{4}$$

We have $\alpha_1^*=17.5$, $\beta_1^*=8.5$, $\alpha_2^*=8.5$, $\beta_2^*=12.5$: H_1 has posterior probability $\Pr(\pi>0.35)=\int_{0.35}^{1}\psi(x;17.5,8.5,8.5,12.5)\mathrm{d}x=0.2855$, and we fail to reject H_0 at the 0.05% level. This means that data combined with our judgment is not enough to make us accept that the difference of these proportions exceeds 0.35. Naturally, different informative, or non-informative, priors can be considered for π_1 and π_2 separately, and the test can be carried out in the same way.

b) Point-null hypothesis:

The point null hypothesis $H_0:\pi=\eta$ vs. $H_1:\pi\ne\eta$ to be tested at the significance level α in Bayesian statistics has been a subject of study and discussion

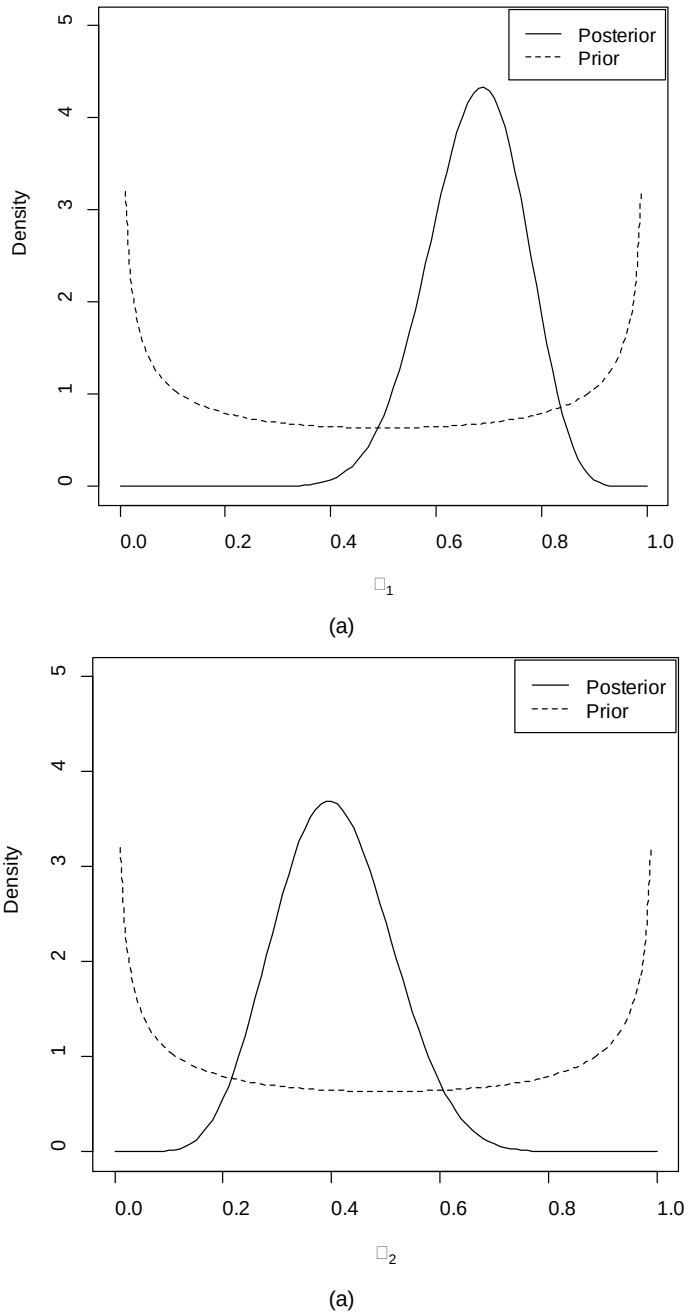


Figure 1. (a) Prior $\text{Beta}(0.5,0.5)$ and posterior $\text{Beta}(17.5,8.5)$ of π_1 and (b) Prior $\text{Beta}(0.5,0.5)$ and posterior $\text{Beta}(8.5,12.5)$ of π_2.

in the literature. Several difficulties still remain concerning this case, especially on the prior probability assigned to the value η (see Berger [10]). We use here Lindley's compromise (Lee [11]), which consists of computing the $(1-\alpha)100\%$ highest posterior density interval and accept or reject H_0 depending on whether η belongs or not to that interval. Here, for the same example, if $\eta = 0.35$, using Pham-Gia and Turkkan's algorithm [12], the 95% hpd interval for π is $(-0.0079; 0.5381)$, which leads us to technically accept H_0 (see **Figure 2**), al-

though the lower bound of the hpd interval can be considered as zero and we can practically reject H_0.

We can see that the above conclusions on π are consistent with each other.

3.2. Bayesian Testing Using the Bayes Factor

Bayesian hypothesis testing can also be carried out using the Bayes factor *B*, which would give the relative weight of the null hypothesis *w.r.t.* the alternative one, when data is taken into consideration. This factor is defined as the ratio of the *posterior odds* over the *prior odds*. With the above expression of the difference of two betas given by (1) we can now accurately compute the Bayes factor associated with the difference of two proportions. We consider two cases:

a) Simple hypothesis: $H_0 : \pi = a$ vs $H_1 : \pi = b$. Then $B = \dfrac{p_\pi(\pi|a)}{p_\pi(\pi|b)}$, which corresponds to the value of the posterior density of π at a, divided by the value of posterior density of π at b. As an application, let us consider the following hypotheses (different from the previous numerical example):

$H_0 : \pi = 0.35$ vs. $H_1 : \pi = 0.25$, where we have uniform priors for both π_1 and π_2, and where we consider the sampling results from Table 1. We obtain the posterior parameters $\alpha_1^* = 18,\ \beta_1^* = 9,\ \alpha_2^* = 9,\ \beta_2^* = 13$. Using the density of the difference (1), we calculate the Bayes factor,

$B = \dfrac{\psi\left(0.35|\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*\right)}{\psi\left(0.25|\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*\right)} = 0.8416$. This value indicates that the data slightly favor H_1 over H_0, which is a logical conclusion since $\hat{\pi} = 0.28$.

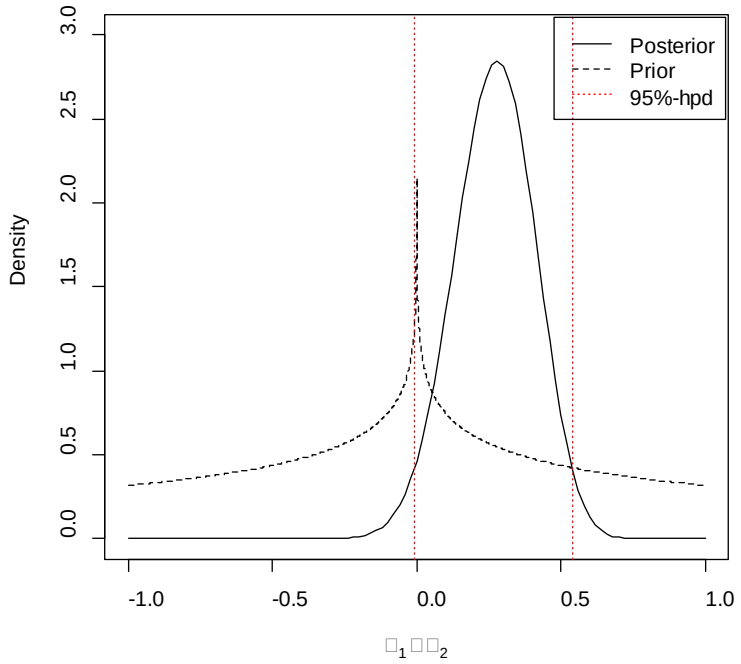


Figure 2. Prior $\psi(0.5, 0.5, 0.5, 0.5)$ and posterior $\psi(17.5, 8.5, 8.5, 12.5)$ distributions of π. The red dashed lines correspond to the bounds of the posterior 95%-hpd interval.

Table 1. Data on customers in area *A* and *B*.

		Area		
		A	*B*	Combined Response
Response	Yes	17	8	25
	No	8	12	20
	Totals	25	20	45

b) Composite hypothesis: As an application, let us consider the hypotheses (4), that is, $H_0 : \pi \le 0.35$ vs. $H_1 : \pi > 0.35$.

In general, $H_0 : \pi \in \Theta_0$ vs. $H_1 : \pi \in \Theta_1$, where $\Theta_0 \cup \Theta_1 = R$. We have $p_0 = \Pr(\pi \in \Theta_0 | \text{posterior})$ and $p_1 = \Pr(\pi \in \Theta_1 | \text{posterior})$ (or $p_1 = 1 - p_0$) as posterior probabilities. Consequently, we define the *posterior odds* on H_0 against H_1 as p_0 / p_1. Similarly, we have the *prior odds* on H_0 against H_1, which we define here as z_0 / z_1. The Bayes factor is $B = \dfrac{p_0 z_1}{p_1 z_0}$. Again, we use the sampling results from Table 1, yielding the prior and posterior distributions presented in Figure 1 with $\text{Beta}(0.5, 0.5)$ prior separately for both proportions.

Now, using (4), $\pi \sim \psi(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)$, we can determine the required prior and posterior probabilities. For example, $p_0 = \int_{-1}^{0.35} \psi(t | \alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*) dt$ gives $p_0 = 0.7145$. In the same way, we obtain $z_0 = 0.745$, using the prior $\psi(1/2, 1/2, 1/2, 1/2)$. Since $p_1 = 1 - p_0$ and $z_1 = 1 - z_0$, we have $p_1 = 0.2855$ and $z_1 = 0.255$. Finally, the Bayes factor is $B = 0.8566$, which is a mild argument in favor of H_1.

4. Prior and Posterior Densities of $\pi - \eta$

The testing above can be seen to be quite straightforward, and is limited to some numerical values of the function $\psi(.)$ that can be numerically computed. But to make an in-depth study of the Bayesian approach to the difference $\pi - \eta = \pi_1 - (\pi_2 + \eta)$, we need to consider the analytic expressions of the prior and posterior distributions of this variable, which can be obtained only from the general beta distribution. Naturally, the related mathematical formulas become more complicated. But Pham-Gia and Turkkan [13] have also established the expression of the density of $X_1 + X_2$, where both have general beta distributions.

4.1. The Difference of Two General Betas

The general beta (or GB), defined on a finite interval, say (*c*, *d*), has a density:

$$f_{gb}(x; \alpha, \beta; c, d) = (x - c)^{\alpha - 1} (d - x)^{\beta - 1} \Big/ \left[(d - c)^{\alpha + \beta - 1} B(\alpha, \beta) \right], \quad \alpha, \beta > 0, \quad c \le x \le d \quad (5)$$

and is denoted by $X \sim GB(\alpha, \beta; c, d)$. It reduces to the standard beta above when $c = 0$ and $d = 1$. Conversely a standard beta can be transformed into a

general beta by addition of, or/and, multiplication with a constant.

Theorem 2: Let $X \sim GB(\alpha,\beta;a,b)$ and any two scalars θ, λ. Then

1) $X+\theta \sim GB(\alpha,\beta;a+\theta,b+\theta)$,

2) $\lambda X \sim GB(\alpha,\beta;\lambda a,\lambda b)$ when $\lambda>0$. Otherwise, $\lambda X \sim GB(\beta,\alpha;\lambda b,\lambda a)$ when $\lambda<0$.

Proof:

1) We have

$$f_{X+\theta}(y) = f_X(y-\theta)$$
$$= \left((y-\theta)-a\right)^{\alpha-1}\left(b-(y-\theta)\right)^{\beta-1} \Big/ \left[(b-a)^{\alpha+\beta-1} B(\alpha,\beta)\right],$$
$$a \le y-\theta \le b$$
$$= \left(y-(a+\theta)\right)^{\alpha-1}\left((b+\theta)-y\right)^{\beta-1} \Big/ \left[\left((b+\theta)-(a+\theta)\right)^{\alpha+\beta-1} B(\alpha,\beta)\right],$$
$$a+\theta \le y \le b+\theta$$

2) For $\lambda>0$,

$$f_{\lambda X}(y) = \frac{1}{\lambda} f_X(y/\lambda)$$
$$= \frac{1}{\lambda}(y/\lambda-a)^{\alpha-1}(b-y/\lambda)^{\beta-1} \Big/ \left[(b-a)^{\alpha+\beta-1} B(\alpha,\beta)\right], a \le y/\lambda \le b$$
$$= (y-\lambda a)^{\alpha-1}(\lambda b-y)^{\beta-1} \Big/ \left[(\lambda b-\lambda a)^{\alpha+\beta-1} B(\alpha,\beta)\right], \lambda a \le y \le \lambda b$$

When $\lambda<0$,

$$f_{\lambda X}(y) = -\frac{1}{\lambda} f_X(y/\lambda)$$
$$= -\frac{1}{\lambda}(y/\lambda-a)^{\alpha-1}(b-y/\lambda)^{\beta-1} \Big/ \left[(b-a)^{\alpha+\beta-1} B(\alpha,\beta)\right], a \le y/\lambda \le b$$
$$= (y-\lambda b)^{\beta-1}(\lambda a-y)^{\alpha-1} \Big/ \left[(\lambda a-\lambda b)^{\alpha+\beta-1} B(\alpha,\beta)\right], \lambda b \le y \le \lambda a$$

Q.E.D.

Pham-Gia and Turkkan [13] gave the expression of the density of X_1+X_2, where X_1 and X_2 are independent general beta variables. The density of X_1-X_2, which is only mentioned there, is explicitly given below.

Proposition 2:

Let $X_1 \sim GB(\alpha,\beta;c,d)$ and $X_2 \sim GB(\gamma,\delta;e,f)$. For the difference X_1-X_2 defined on $(c-f,d-e)$, there are two different cases to consider, depending on the relative values of $c-e$ and $d-f$, since X_1 and X_2 do not have symmetrical roles.

Case 1:

$$c-f \le d-f \le c-e \le d-e \tag{6}$$

Case 2:

$$c-f \le c-e \le d-f \le d-e \tag{7}$$

Theorem 3: Let X_1 and X_2 be two independent general betas with their supports satisfying (6). Then $Y=X_1-X_2$ has its density defined as follows:

For $c-f \le y \le d-f$,

$$f(y)=\frac{(y-(c-f))^{\alpha+\delta-1}(d-f-y)^{\beta-1}B(\delta,\alpha)}{(d-c)^{\alpha+\beta-1}(f-e)^{\delta}B(\delta,\gamma)B(\alpha,\beta)} F_1\left(\delta,1-\beta,1-\gamma;\alpha+\delta;\frac{(c-f)-y}{(d-f)-y},\frac{y-(c-f)}{f-e}\right) \quad (8)$$

For $d-f\le y\le c-e$,

$$f(y)=\frac{(y-(d+f))^{\delta-1}(d-e-y)^{\gamma-1}}{(f-e)^{\delta+\gamma-1}B(\delta,\gamma)} F_1\left(\beta,1-\delta,1-\gamma;\alpha+\beta;\frac{c-d}{y-(d-f)},\frac{d-c}{d-e-y}\right) \quad (9)$$

and for $c-e\le y\le d-e$,

$$f(y)=\frac{((d-e)-y)^{\beta+\gamma-1}(y-(d-f))^{\delta-1}B(\beta,\gamma)}{(d-c)^{\beta}(f-e)^{\delta+\gamma-1}B(\delta,\gamma)B(\alpha,\beta)} F_1\left(\beta,1-\alpha,1-\delta;\beta+\gamma;\frac{(d-e)-y}{d-c},\frac{y-(d-e)}{y-(d-f)}\right) \quad (10)$$

where $F_1(.)$ is Appell's first hypergeometric function already discussed.

Proof:

The argument uses first part 2) of **Theorem 1** to obtain that $-X_2\sim GB(\delta,\gamma;-f,-e)$. Then, it uses the exact expression of the density of the sum of two general betas (see **Theorem 2** in the article of T. Pham-Gia & N. Turkkan [14]).

Q.E.D.

We denote the above density given by (8), (9) and (10) by $\varphi_{\pi}(\alpha_1,\beta_1,\alpha_2,\beta_2;c,d,e,f)$

Note: The corresponding case 2, when relation (7) is satisfied, is given in **Appendix 1** (**Theorem 3a**).

To study the density of $\pi-\eta=\pi_1-(\pi_2+\eta)$, a particular case that will be used in our study here is the difference between $X_1\sim GB(\alpha_1,\beta_1;0,1)$ and $X_2\sim GB(\alpha_2,\beta_2;\eta,\eta+1),-1\le\eta\le 1$, with η being a positive constant.

In this case both **Theorem 2** and **Theorem 3** apply since $c-e=d-f$ and the middle definition section of $\varphi_{\pi}(\alpha_1,\beta_1,\alpha_2,\beta_2;c,d,e,f)$ disappears.

Theorem 4: Let $X_1\sim GB(\alpha_1,\beta_1;0,1)$ and $X_2\sim GB(\alpha_2,\beta_2;\eta,\eta+1)$ be two independent general beta distributed random variables. Then the density of $Y=X_1-X_2$, defined on $[-(\eta+1),1-\eta]$, is:

1) for $-\eta-1\le y\le-\eta$,

$$f(y)=\frac{(y+(\eta+1))^{\alpha_1+\beta_2-1}(-\eta-y)^{\alpha_2-1}B(\alpha_1,\beta_2)}{B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)} F_1\left(\beta_2,1-\beta_1,1-\alpha_2;\alpha_1+\beta_2;\frac{(\eta+1)+y}{\eta+y},y+(\eta+1)\right)$$

2) for $-\eta\le y\le 1-\eta$,

$$f(y)=\frac{\left((1-\eta)-y\right)^{\alpha_2+\beta_1-1}(y+\eta)^{\beta_2-1}B(\alpha_2,\beta_1)}{B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)}$$
$$F_1\left(\beta_1,1-\alpha_1,1-\beta_2;\alpha_2+\beta_1;(1-\eta)-y,\frac{y-(1-\eta)}{y+\eta}\right)$$

and we denote this distribution by

$$Y\sim\xi_\eta(\alpha_1,\beta_1,\alpha_2,\beta_2;\eta). \tag{11}$$

Proof:

This is a special case of **Theorem 3**.

Q.E.D.

An equivalent form using **Theorem 4** leads to a slightly different expression, which gives however, the same numerical values for the density of $\pi-\eta$ (see **Theorem 4a** in Appendix 1).

4.2. Prior and Posterior Distributions of $\pi-\eta$

Let π_i, $i=1,2$ be two independent beta distributed random variables, the first being a regular beta, $\pi_1\sim\text{beta}(\alpha_1,\beta_1)$, and the second being a general beta, $\pi_2\sim GB(\alpha_2,\beta_2;\eta,1+\eta)$.

Binomial sampling, with these two different beta priors, leads to the following

Proposition 3: The prior distribution of $\pi-\eta=\pi_1-(\pi_2+\eta)$ is $\xi_\eta(\alpha_1,\beta_1,\alpha_2,\beta_2;\eta)$, given by (11), and its posterior distribution is $\xi_\eta(\alpha_1^*,\beta_1^*,\alpha_2^*,\beta_2^*;\eta)$ with $\alpha_i^*=\alpha_i+x_i$ and $\beta_i^*=\beta_i+n_i-x_i$, $i=1,2$.

Proof:

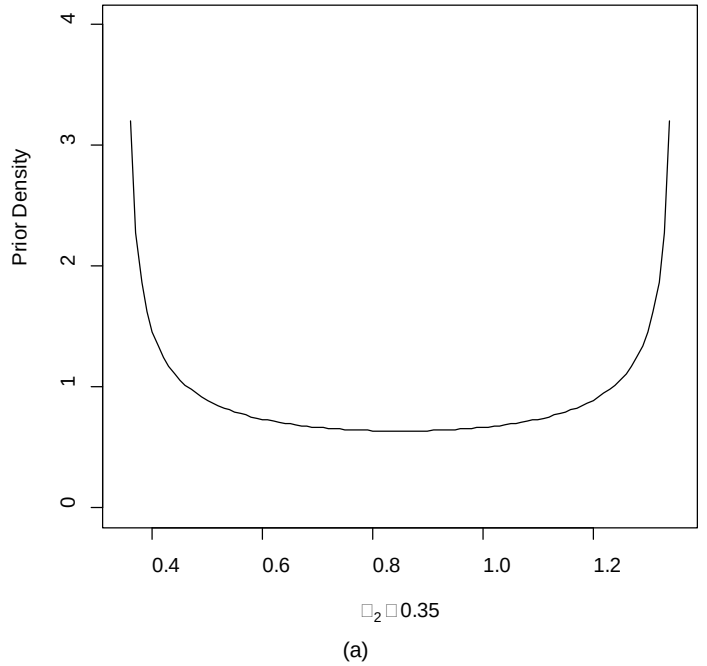
$\pi_1-(\pi_2+\eta)$ is the difference of two random variables with respective distribution $\text{beta}(\alpha_1,\beta_1)$ and $GB(\alpha_2,\beta_2;\eta,\eta+1)$, The prior distributions of $\pi-\eta$ is hence $\xi_\eta(\alpha_1,\beta_1,\alpha_2,\beta_2;\eta)$, as given by (14).

Binomial sampling affects these 2 distributions in different ways. For the first, the posterior is $\text{beta}(\alpha_1+x_1,\beta_1+n_1-x_1)$ while the posterior distribution of the second is $GB(\alpha_2+x_2,\beta_2+n_2-x_2;\eta,\eta+1)$ (see **Proposition 3a** in Appendix 2). Figure 3 shows the prior and the posterior of $\pi_2+0.35$.

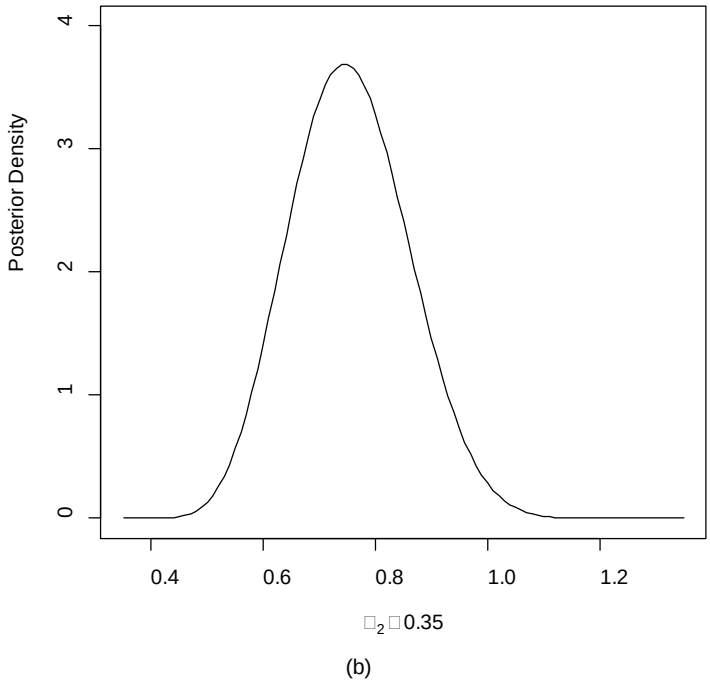
From **Theorem 4**, we obtain the expression of the posterior density $\xi_{.35}(17.5,8.5,8.5,12.5;0.35)$ of $\pi-\eta$ as follows:

$$f(x)=\begin{cases}\dfrac{(x+1.35)^{29}(-0.35-x)^{7.5}B(17.5,12.5)}{B(17.5,8.5)B(8.5,12.5)} & \\ \quad F_1\left(12.5,-7.5,-7.5;30;\dfrac{1.35+x}{0.35+x},x+1.35\right), & \\ \qquad -1.35\le x<-0.35 & \\ \dfrac{(0.65-x)^{16}(x+0.35)^{11.5}B(8.5,8.5)}{B(17.5,8.5)B(8.5,12.5)} & \\ \quad F_1\left(8.5,-16.5,-11.5;17;0.65-x,\dfrac{x-0.65}{x+0.35}\right), & \\ \qquad -0.35\le x<0.65 & \end{cases} \tag{12}$$

Figure 4 shows the above density.



(a)



(b)

Figure 3. (a) Prior $GB(0.5,0.5,0.35,1.35)$ distribution of $\pi_2+0.35$ and (b) Posterior $GB(8.5,12.5;0.35,1.35)$ distribution of $\pi_2+0.35$. The posterior of $\pi_1-(\pi_2+\eta)$ is hence given by **Theorem 4**, as $\xi_\eta(\alpha_1^*,\beta_1^*,\alpha_2^*,\beta_2^*;\eta)$.

5. Conclusion

The Bayesian approach to testing the difference of two independent proportions leads to interesting results which agree with frequentist results when non-informative priors are considered. Undoubtedly, all preceding results can be

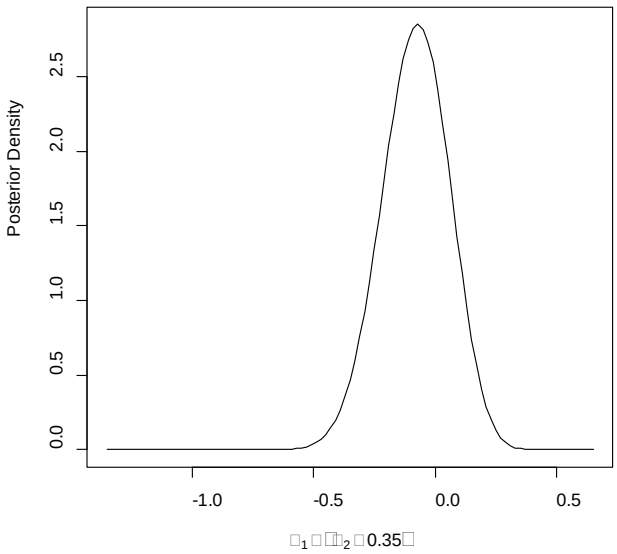


Figure 4. Posterior density $\xi_{.35}(17.5,8.5,8.5,12.5;0.35)$ of $\pi_1-(\pi_2+0.35)$.

generalized to other measures frequently used in a 2 × 2 table.

Acknowledgements

Research partially supported by NSERC grant 9249 (Canada). The authors wish to thank the Universite de Moncton Faculty of Graduate Studies and Research for the assistance provided while conducting this work.

References

[1] Pham-Gia, T. and Turkkan, N. (1993) Bayesian Analysis of the Difference of two Proportions. *Communications in Statistics— Theory and Methods*, **22**, 1755-1771. https://doi.org/10.1080/03610929308831114

[2] Pham-Gia, T. and Turkkan, N. (2008) Bayesian Analysis of a 2 × 2 Contingency Table with Dependent Proportions and Exact Sample Sizes. *Statistics*, **42**, 127-147. https://doi.org/10.1080/02331880701600380

[3] Pham-Gia, T. and Turkkan, N. (2003) Determination of the Exact Sample Sizes in the Bayesian Estimation of the Difference between Two Proportions. *Journal of the Royal Statistical Society*, **52**, 131-150. https://doi.org/10.1111/1467-9884.00347

[4] D'Agostino, R., Chase, W. and Belanger, A. (1988) The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations. *The American Statistician*, **42**, 198-202.

[5] Cressie, N. (1978) Testing the Equality of Two Binomial Proportions. *Annals of the Institute of Statistical Mathematics*, **30**, 421-427. https://doi.org/10.1007/BF02480232

[6] Eberhardt, K.R. and Fligner, M.A. (1977) A Comparison of Two Tests for Equality of Two Proportions. *The American Statistician*, **21**, 151-155. https://doi.org/10.1080/00031305.1977.10479225

[7] Santner, T.J. and Snell, M.K. (1975) Small Sample Confidence Intervals for p_1-p_2 and p_1/p_2 in 2 × 2 Contingency Tables. *JASA*, **75**, 386-394.

[8] Howard, J.V. (1998) The 2 × 2 Table: A Discussion from a Bayesian Viewpoint. *Statistical Sciences*, **13**, 351-367. https://doi.org/10.1214/ss/1028905830

[9] Agresti, A. and Coull, B. (1998) Approximate Is Better than Exact for Interval Estimation of Binomial Proportions. *The American Statistician*, **52**, 2.

[10] Berger, J. (1999) Bayes Factor. In: Kotz, S., Read, C.B. and Banks, D.L., Eds., *Encyclopedia of Statistics*, Update 3, Wiley, NY, 20-29.

[11] Lee, P.M. (2004) Bayesian Statistics. An Introduction. 3rd Edition, Hodder Arnold, London.

[12] Pham-Gia, T. and Turkkan, N. (1993) Computation of the Highest Posterior Density Interval in Bayesian Analysis. *Journal of Statistical Computation and Simulation*, **44**, 243-250. https://doi.org/10.1080/00949659308811461

[13] Pham-Gia, T. and Turkkan, N. (1998) Distribution of the Linear Combination of Two General Beta Variables and Applications. *Communications in Statistics—Theory and Methods*, **27**, 1851-1869. https://doi.org/10.1080/03610929808832194

[14] Pham-Gia, T. and Turkkan, N. (1994) Reliability of a Standby System with Beta-Distributed Component Lives. *IEEE Transactions on Reliability*, **R43**, 71-75. https://doi.org/10.1109/24.285114

Appendix 1

Below is the expression of the density of $Y = X_1 - X_2$ when (7) is satisfied, instead of (6). This expression, with the one given in **Theorem 3**, covers all cases.

Theorem 3a: Let X_1 and X_2 be two independent general betas with their supports satisfying (10). Then $Y = X_1 - X_2$ has its density defined as follows: for $c - f \le y \le c - e$,

$$f(y) = \frac{\left(y-(c-f)\right)^{\alpha+\delta-1}(c-e-y)^{\gamma-1}B(\alpha,\delta)}{(f-e)^{\delta+\gamma-1}(d-c)^{\alpha}B(\alpha,\beta)B(\delta,\gamma)} F_1\left(\alpha,1-\gamma,1-\beta;\alpha+\delta;\frac{(c-f)-y}{(c-e)-y},\frac{y-(c-f)}{d-c}\right) \tag{13}$$

For $c - e \le y \le d - f$,

$$f(y) = \frac{\left(y+(c+e)\right)^{\alpha-1}(d-e-y)^{\beta-1}}{(d-c)^{\alpha+\beta-1}B(\alpha,\beta)} F_1\left(\gamma,1-\alpha,1-\beta;\delta+\gamma;\frac{e-f}{y-(c-e)},\frac{f-e}{d-e-y}\right) \tag{14}$$

For $d - f \le y \le d - e$,

$$f(y) = \frac{\left((d-e)-y\right)^{\beta+\gamma-1}\left(y-(c-e)\right)^{\alpha-1}B(\beta,\gamma)}{(f-e)^{\gamma}(d-c)^{\alpha+\beta-1}B(\delta,\gamma)B(\alpha,\beta)} F_1\left(\gamma,1-\delta,1-\alpha;\beta+\gamma;\frac{(d-e)-y}{f-e},\frac{y-(d-e)}{y}\right) \tag{15}$$

Proof:

By rewriting $Y = (-X_2) - (-X_1)$, we can apply the above **Theorem 2** and **Theorem 3**.

Q.E.D

A parallel, and equivalent, result to **Theorem 4** is given below:

Theorem 4a: The density of $X_1 - X_2 - \eta$ is:

For $-\eta - 1 \le y \le -\eta$,

$$f(y) = \frac{\left(y+(\eta+1)\right)^{\alpha_1+\beta_2-1}(-\eta-y)^{\alpha_2-1}B(\alpha_1,\beta_2)}{B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)} F_1\left(\alpha_1,1-\alpha_2,1-\beta_1;\alpha_1+\beta_2;\frac{(\eta+1)+y}{\eta+y},y+(\eta+1)\right)$$

For $-\eta \le y \le 1 - \eta$,

$$f(y) = \frac{\left((1-\mu)-y\right)^{\alpha_2+\beta_1-1}(y+\eta)^{\alpha_1-1}B(\alpha_2,\beta_1)}{B(\alpha_1,\beta_1)B(\alpha_2,\beta_2)} F_1\left(\alpha_2,1-\beta_2,1-\alpha_1;\alpha_2+\beta_1;(1-\eta)-y,\frac{y-(1-\eta)}{y+\eta}\right)$$

and we denote $Y \sim \xi_{\eta}^{*}(\alpha_1,\beta_1,\alpha_2,\beta_2;\eta)$.

Proof:

Similar to the proof of **Theorem 4**.

Q.E.D

Appendix 2

Proposition 3a:

Suppose that $X_2 \sim \text{Bin}(n_2, \pi_2)$ and π_2 has the prior distribution beta(α_2, β_2) then the posterior distribution of $\pi_2 + \eta$ is $GB(\alpha_2 + x_2, \beta_2 + n_2 - x_2; \eta, \eta + 1)$.

Proof:

The prior distribution of $\pi_2 + \eta$ is $GB(\alpha_2, \beta_2; \eta, \eta + 1)$ (see **Theorem 2**) with the pdf

$$f_{\pi_2+\eta}(\pi_2|x_2) = \left[B(\alpha_2, \beta_2)\right]^{-1}(\pi_2 - \eta)^{\alpha_2 - 1}(1 + \eta - \pi_2)^{\beta_2 - 1},\ \eta \le \pi_2 \le \eta + 1,$$

The likelihood function is

$$f_{X_2|\pi_2+\eta}(x_2|\theta) = f_{X_2|\pi_2}(x_2|\pi_2) = \binom{n_2}{x_2}\pi_2^{x_2}(1-\pi_2)^{n_2-x_2},\ x_2 = 0,1,\cdots,n$$

Thus the marginal distribution of X_2, the number of success, with $\pi_2 = \theta - \eta$, has density:

$$\begin{aligned}
K(x_2|\alpha_2, \beta_2, n_2) &= \frac{\binom{n_2}{x_2}}{B(\alpha_2, \beta_2)}\int_\eta^{\eta+1}(\theta-\eta)^{\alpha_2-1}(1+\eta-\theta)^{\beta_2-1}\pi_2^{x_2}(1-\pi_2)^{n_2-x_2}\,\mathrm{d}\theta, \\
&= \frac{\binom{n_2}{x_2}}{B(\alpha_2, \beta_2)}\int_\eta^{\eta+1}(\theta-\eta)^{\alpha_2-1}(1+\eta-\theta)^{\beta_2-1}(\theta-\eta)^{x_2}(1+\eta-\theta)^{n_2-x_2}\,\mathrm{d}\theta \\
&= \frac{\binom{n_2}{x_2}}{B(\alpha_2, \beta_2)}\int_\eta^{\eta+1}(\theta-\eta)^{\alpha_2+x_2-1}(1+\eta-\theta)^{\beta_2+n_2-x_2-1}\,\mathrm{d}\theta \\
&= \frac{\binom{n_2}{x_2}}{B(\alpha_2, \beta_2)}B(\alpha_2 + x_2, \beta_2 + n_2 - x_2)
\end{aligned}$$

Therefore, the posterior distribution of θ given $X_2 = x_2$ is

$$\begin{aligned}
f_{\pi_2+\eta|X_2}(\theta|x_2) &= \frac{f_{\pi_2+\eta}(\theta|x_2)\, f_{X_2|\pi_2+\eta}(x_2|\theta)}{K(x_2|\alpha_2, \beta_2, n_2)} \\
&= \frac{\left[B(\alpha_2, \beta_2)\right]^{-1}(\theta-\eta)^{\alpha_2-1}(1+\eta-\theta)^{\beta_2-1}\binom{n_2}{x_2}\pi_2^{x_2}(1-\pi_2)^{n_2-x_2}}{\frac{\binom{n_2}{x_2}}{B(\alpha_2, \beta_2)}B(\alpha_2 + x_2, \beta_2 + n_2 - x_2)},
\end{aligned}$$

with $\pi_2 = \theta - \eta,\ \eta \le \theta \le \eta + 1$

$$= \frac{(\theta-\eta)^{\alpha_2+x_2-1}(1+\eta-\theta)^{\beta_2+n_2-x_2-1}}{B(\alpha_2 + x_2, \beta_2 + n_2 - x_2)}, \eta \le \theta \le \eta + 1$$

This is the p.d.f. of $GB(\alpha_2 + x_2, \beta_2 + n_2 - x_2; \eta, \eta + 1)$.

Q. E. D.

End