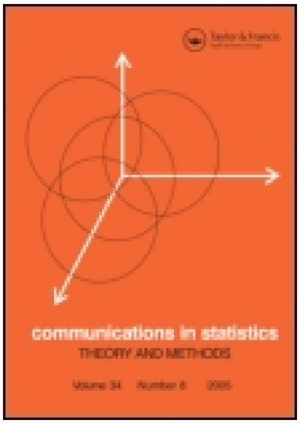

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Bayesian analysis of the difference of two proportions

T. Pham-Gia Ph.D.,FIS [a] , N. Turkkan [b] & P. Eng [b]
[a] Faculty of Science, Université de Moncton, Moncton, N.B., E1A 3E9
[b] School of Engineering, Université de Moncton, Moncton, N.B., E1A 3E9
Published online: 27 Jun 2007.

COMMUN. STATIST.—THEORY METH., 22(6), 1755-1771 (1993)

BAYESIAN ANALYSIS OF THE DIFFERENCE OF TWO PROPORTIONS

T. Pham-Gia, Ph.D., FIS,

Faculty of Science
Université de Moncton
Moncton, N. B.
ElA 3E9

N. Turkkan, P. Eng.

School of Engineering
Université de Moncton
Moncton, N. B.
ElA 3E9

Key Words & Phrases: proportions; beta distribution; Appell function; beta–binomial; non informative prior

ABSTRACT

In this article we give the expression of the prior distribution of $p_1 - p_2$, where p_1 and p_2 are independent proportions with a beta prior each. The expression derived for the posterior distribution of $p_1 - p_2$ then shows the closure of the beta–difference family for independent dual Bernoulli samples. Other bayesian results are also presented.

1. INTRODUCTION

One of the basic results in bayesian analysis is the natural conjugate property of the beta family for binomial sampling: if p has a beta (α, β) prior then the posterior distribution of p is beta $(\alpha + x, \beta + n - x)$, where

x and n are respectively the number of successes and the total number of observations in the sampling phase.

For two independent binomial parameters p_1 and p_2, the difference $p = p_1 - p_2$, with each p_i having a beta (α_i, β_i) prior, $i = 1, 2$, is sometimes studied in a bayesian context. Such a difference has been considered, for example, by Wright (1988) for sample allocation purposes and by Draper and Guttman (1969) for information ratio computation. The lack of further results concerning the bayesian treatment of $p_1 - p_2$ mostly stems from the fact that the precise prior density of $p_1 - p_2$, i.e. the difference of two beta's, is unknown. If this density can be established it can be used in several important applications, especially in the area of bayesian optimal allocation, along the same line as Zacks's work (1970). The current practical approach, however, consists in finding another beta which would adequately approximate it, by matching either moments (Springer (1979, p. 268)) or percentiles.

Here, we will first establish the precise expression of that density and then apply it to compute the posterior distribution. Other results presented concern the determination of the highest posterior density (hpd) region, the predictive distribution of $X_1 - X_2$ and the Bayes risk associated with a squared–error loss function. A numerical example illustrates the above results.

2. The difference of two betas.

Theoretically, the density of the sum or difference of 2 independent random variables can be obtained by Fourier transform methods (Springer

(1979, p. 52)). However, in practice, difficulties quickly arise when trying to obtain the precise expression of this density. This is the case of the beta distribution and, to our knowledge, only Irwin (1927) gave precise, but very elaborate, formulas for the sum of independent and identical beta's.

The connection between the beta (and its generalizations) and hypergeometric functions has been presented in Pham–Gia and Duong (1989). The same connection will permit us here to obtain the desired density. There are four Appell hypergeometric functions in two variables, F_1, F_2, F_3 and F_4. Only F_1 and F_3, which are closely related to each other, will be used here. We also recall a fundamental theorem of Picard which relates Appell's first hypergeometric function in 2 variables to an eulerian integral in a single variable.

Definition: Let a_1, a_2, b_1, b_2 and c be real or complex numbers with c different from a negative integer. Appell's first hypergeometric function in 2 variables, F_1, is defined by:

$$F_1(a_1, b_1, b_2; c; x_1, x_2) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(a_1, m+n)}{(c, m+n)} (b_1, m)(b_2, n) \frac{x_1^m}{m!} \frac{x_2^n}{n!} ,$$

where $(a, m) = a(a+1) \ldots (a+m-1) = \Gamma(a+m)/\Gamma(a)$, $m > 0$, with $(a, 0) = 1$. Similarly, F_3 is defined by:

$$F_3(a_1, a_2, b_1, b_2; c; x_1, x_2) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(a_1, m)(a_2, n)}{(c, m+n)} (b_1, m)(b_2, n) \frac{x_1^m}{m!} \frac{x_2^n}{n!} .$$

It is established that F_1 and F_3 converge for $|x_1| < 1$ and $|x_2| < 1$.

Appell functions are generalizations to 2 variables of the well–known Gauss hypergeometric function $_2F_1(a, b; c; x) = \sum_{m=0}^{\infty} \frac{(a,m)}{(c,m)} (b,m) \frac{x^m}{m!}$.

Let's recall that ${}_2F_1$ contains $\exp(x)$ as a special case and has an integral representation of the form

$${}_2F_1(a, b; c; x) = \int_0^1 u^{a-1} (1-u)^{c-a-1} (a - xu)^{-b} du/B(a, c-a)$$

Extending this representation to 2 variables, Picard obtained:

Theorem (Picard): Let a, b_1, b_2 and c be real or complex numbers. If $\mathrm{Re}(a)$ and $\mathrm{Re}(c-a)$ are positive and $F_1(a, b_1, b_2; c; x_1, x_2)$ converges, then

$$F_1(a, b_1, b_2; c; x_1, x_2) = \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(c-a)} \int_0^1 u^{a-1} (1-u)^{c-a-1} (1-ux_1)^{-b_1} \cdot (1-ux_2)^{-b_2}\, du$$

Proof: See (Exton (1976, p. 27)). No similar integral representation has been found for the remaining Appell functions.

In the more general context of hypergeometric functions in n variables F_1 and F_3 are special cases, when $n = 2$, of the Lauricella D – functions $F_D^{(n)}$ and B – functions $F_B^{(n)}$ (see Pham–Gia and Duong (1989)).

Among numerous properties of Appell's hypergeometric functions we will retain the following ones related to changes of variables in F_1:

$$F_1(a, b_1, b_2; c; x, y) = (1-x)^{c-(a+b_1)}(1-y)^{-b_2} \cdot F_1(c-a, c-(b_1+b_2), b_2; c; x, (y-x)/(y-1)) \tag{1a}$$

$$= (1-x)^{-b_1}(1-y)^{c-(a+b_2)}\, F_1(c-a, b_1, c-(b_1+b_2); c; (x-y)/(x-1), y) \tag{1b}$$

Finally, F_1 and F_3 are related by the following relations:

$$F_1(a, b_1, b_2; c; x, y) = (1-y)^{-b_2} F_3(a, c-a, b_1, b_2; c; x, y/(y-1)) \tag{1c}$$

$$= (1-x)^{-b_1} F_3(c-a, a, b_1, b_2; c; x/(x-1), y) \tag{1d}$$

The proofs of the above equalities are given in Appell and Kampé de Fériet (1926, p. 30). We can now establish the following:

Theorem: Let $p_i \sim \text{beta}(\alpha_i, \beta_i)$, $i = 1, 2$, be independent variables. Then $p = p_1 - p_2$ has the following density:

For $0 < p \le 1$, $f(p) =$

$$B(\alpha_2, \beta_1)\, p^{\beta_1+\beta_2-1} (1-p)^{\alpha_2+\beta_1-1} F_1(\beta_1, \alpha_1+\beta_1+\alpha_2+\beta_2-2, 1-\alpha_1; \beta_1+\alpha_2; (1-p), 1-p^2)/A \tag{2a}$$

and for $-1 \le p < 0$

$$f(p) = B(\alpha_1, \beta_2)\, (-p)^{\beta_1+\beta_2-1} (1+p)^{\alpha_1+\beta_2-1} F_1(\beta_2, 1-\alpha_2, \alpha_1+\alpha_2+\beta_1+\beta_2-2; \alpha_1+\beta_2; 1-p^2, 1+p)/A, \tag{2b}$$

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

Moreover, if $\alpha_1 + \alpha_2 > 1$ and $\beta_1 + \beta_2 > 1$, we have:

$$f(0) = B(\alpha_1 + \alpha_2 - 1, \beta_1 + \beta_2 - 1)/A. \tag{2c}$$

Proof (summary): Since $f(p)$ is the convolution of f_1 and f_2, $f = f_1 * f_2$, for $0 < p \le 1$ we have:

$$f(p) = \int_0^{1-p} (p+v)^{\alpha_1-1} (1-p-v)^{\beta_1-1} v^{\alpha_2-1} (1-v)^{\beta_2-1}\, dv/A \tag{3}$$

$$= p^{\alpha_1-1} (1-p)^{\beta_1-1} \int_0^{1-p} v^{\alpha_2-1} (1-v)^{\beta_2-1} \left(1+\frac{v}{p}\right)^{\alpha_1-1} \cdot \left(1-\frac{v}{1-p}\right)^{\beta_1-1} dv/A,$$

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

Changing the integration variable to $w = v/(1-p)$ and applying Picard's Theorem, we obtain:

$$f(p) = p^{\alpha_1-1}(1-p)^{\alpha_2+\beta_1-1} B(\alpha_2, \beta_1) F_1(\alpha_2, 1-\beta_2, 1-\alpha_1; \beta_1+\alpha_2; 1-p, 1-\tfrac{1}{p})/A. \tag{4}$$

Applying equation (1a) to change the variables inside F_1, we obtain equation (2a) where $|1-p| < 1$ and $|1-p^2| < 1$, as required for the convergence of F_1. Equation (2b) can be proved in a similar way.

For $p = 0$ equation (3) above reduces immediately to (2c) if $\alpha_1 + \alpha_2 > 1$ and $\beta_1 + \beta_2 > 1$. However, $f(0)$ can be undefined in other cases.

QED

<u>Remarks</u>: 1) Since beta (α, β) can have a wide variety of shapes depending on the values of α and β, the density of $p = p_1 - p_2$ will also have a wide variety of shapes. For example, in equation (2a) if $\beta_1 + \beta_2 \geq 1$ and F_1 is finite then $f(p)$ converges to a finite limit when $p \to 0+$ but diverges to ∞ if $\beta_1 + \beta_2 < 1$. A similar remark holds for $\alpha_2 + \beta_1$ when $p \to 1-$ in the same equation. For equation (2b) a discussion similar to the above can be made.

2) When $\alpha_1 \geq 1$, applying relation (1d) to equation (4), $f(p)$ can be expressed in terms of an F_3 function and is slightly more convenient to use since $f(0)$ can now be computed directly from the expression obtained:

$$f(p) = (1-p)^{\alpha_2+\beta_1-1} B(\alpha_2, \beta_1) F_3(\beta_1, \alpha_2, 1-\alpha_1, 1-\beta_2; \alpha_2+\beta_1; 1-p, 1-p)/A$$

for $0 \leq p \leq 1$.

A similar expression holds on $-1 < p \leq 0$ if $\alpha_2 \geq 1$.

3) Since $f_1 * f_2 = f_2 * f_1$, an equivalent form for $f(p)$ can be obtained by changing α_1 to α_2 and β_1 to β_2 (and vice versa) in equations (2a), (2b) and (2c).

3. The posterior distribution

3.1 When independent sampling of p_1 and p_2 gives respectively x_1 and x_2 favorable outcomes out of n_1 and n_2 observations, the posterior distributions of p_1 and p_2 are then independent beta $(\alpha_1 + x_1, \beta_1 + n_1 - x_1)$ and beta $(\alpha_2 + x_2, \beta_2 + n_2 - x_2)$. The posterior distribution of p, $f(p \mid n_1, x_1, n_2, x_2)$ is then given by equations (2a), (2b) and (2c), with $\alpha_i + x_i$ replacing α_i and $\beta_i + (n_i - x_i)$ replacing β_i, $i = 1, 2$. This property leads to the following definition and proposition.

Definition: Let α_1, β_1, α_2 and β_2 be positive numbers. A distribution with its density given by equations (2a), (2b) and (2c) is called a beta–difference distribution.

As stated in remark 1 of the previous section, the beta–difference family can have a wide variety of shapes depending on the values of α_i and β_i. Graphs in sections 4 and 5 illustrate some of these shapes.

Proposition: The beta–difference distribution family is closed under independent dual Bernoulli sampling.

3.2 The precise expression of the posterior also allows the computation of the highest posterior density (hpd) region for p. Without that expression,

the approximate $100(1-\alpha)\%$ credible interval for p is usually computed, in practice, by using $\mu_{post}(p) \pm c_{1-\alpha/2}\sqrt{var_{post}(p)}$ (5)
(see Berger (1985, p. 136)), where

the posterior mean of p is $\mu_{post}(p) = \dfrac{\alpha_1 + x_1}{\alpha_1 + \beta_1 + n_1} - \dfrac{\alpha_2 + x_2}{\alpha_2 + \beta_2 + n_2}$,

the posterior variance is

$$Var_{post}(p) = \sum_{i=1}^{2} \frac{(\alpha_i + x_i)(\beta_i + n_i - x_i)}{(\alpha_i + \beta_i + n_i)^2(\alpha_i + \beta_i + n_i + 1)} \tag{6}$$

and $c_{1-\alpha/2}$ is the $1-\alpha/2$ percentile of the normal distribution. The convergence of the posterior toward the normal is usually used as a justification for using $c_{1-\alpha/2}$. Naturally, this practice has a varying degree of success, depending on the rate of convergence of the posterior toward the normal.

The highest posterior density (hpd) credible interval $I_{1-\alpha}$ is often numerically computed although tables exist for some distributions (see Isaacs, Christ, Novick and Jackson (1974)). $I_{1-\alpha}$ has the property that $P(p \in I_{1-\alpha}) = 1-\alpha$ and at the same time $f(y) < f(p)$ for any $y \notin I_{1-\alpha}$ and any $p \in I_{1-\alpha}$. Berger (1985) has proposed an algorithm to determine this region and a computer program developped by Turkkan and Pham–Gia (1992) computes the same interval using a different approach. It is available from the authors upon request. It has been successfully used in a number of applications (see Pham–Gia and Turkkan (1992)). Here, it is applied in the numercial example in the final section.

In bayesian decision theory, the Bayes risk associated with the squared error loss function, $L(p, \hat{p}) = c(p - \hat{p})^2$, where $\hat{p}$ is an estimation of p, is shown to be $c\,E_{x_1,x_2}(Var_{post}(p))$ where Var_{post} is the posterior variance and E_{x_1,x_2} is the expectation with respect to X_1 and X_2 taken as variables.

Due to the independence of p_1 and p_2, we have:

$$E_{x_1,x_2}(\mathrm{Var}_{post}(p)) = E_{x_1}(\mathrm{Var}_{post}(p_1)) + E_{x_2}(\mathrm{Var}_{post}(p_2))$$
$$= \sum_{i=1}^{2} \frac{\alpha_i \ \beta_i}{(\alpha_i+\beta_i)(\alpha_i+\beta_i+1)(\alpha_i+\beta_i+n_i)}.$$

Finally, for n_i fixed, the marginal distribution of X_i, also called predictive distribution, is a beta–binomial distribution with density:

$$P(X_i = k) = \binom{n_i}{k} \frac{B(\alpha_i+k, \beta_i+n_i-k)}{B(\alpha_i, \beta_i)}, \quad k = 0, 1, ..., n_i, i = 1, 2 \tag{7}$$

Using (7) and the rules for the density of the difference of two discrete variables, we can establish that $X = X_1 - X_2$ has as distribution:

$$P(X = k) = \frac{1}{A} \sum_{\ell=m(k)}^{M(k)} \binom{n_1}{\ell} \binom{n_2}{\ell-k} B(\alpha_1 + \ell, \beta_1 + n_1 - \ell)\, B(\alpha_2 + \ell - k, \beta_2 + n_2 - \ell + k), \tag{8}$$

with $k = -n_2, -n_2 + 1, ..., n_1$, $m(k) = \max(0,k)$ and $M(k) = \min(n_1, n_1 + k)$.

Moments of X can be computed directly from (8) but, in most cases, they are more conveniently obtained by applying the binomial formula:

$E(X^n) = \sum_{j=0}^{n} \binom{n}{j} (-1)^{n-j}\, E\left[X_1^j\right] E\left[X_2^{n-j}\right]$, $n = 1, 2, ..., i=1, 2$ where $E\left[X_i^j\right]$, being the moments of the beta–binomial distribution, can be obtained in closed forms.

3.3 Bayesian robustness is a very complex topic on which considerable discussion has been devoted in the literature (for a general survey see Berger (1985, sect. 4.7)). Several approaches to bayesian robustness can be taken, but in this section, we limit ourselves to prior robustness whereby the relationships between changes in the posterior are evaluated against those introduced in the prior. Changes in the posterior can be measured by a variety of criteria including posterior $(1-\alpha)100\%$ credible interval, posterior mean, posterior variance, Bayes risk (see Pham–Gia (1990)). Generally, these criteria vary in the same direction with, however, considerable differences in the variations of their magnitude. The posterior credible interval seems to be the most appropriate criterion to consider and changes in the prior are generally introduced through changes in the values of its hyperparameters.

Basically, since the posterior is proportional to the product of the prior with the likelihood function, the posterior is not too sensible to changes in the prior if the likelihood function, which depends on sampling results, dominates in this product. This domination can be traced to the peakness of the likelihood function in comparison to the region of the prior on which it operates. Also, in practice, if this region is the "body" of the prior "there is less concern for prior robustness because this body is comparatively easy to specify whereas the tail is very hard to specify" (Berger (1985, p. 223)).

Beta distributions are not automatically robust as some studies seem to imply (see Pham–Gia (1990)). However, similarly to other natural conjugate priors, they are robust if the likelihood function is concentrated in their central portion. For the beta–difference family presented here the same conclusions apply, separately for each prior beta. The numerical example in section 5 illustrates this point.

4. Non–informative priors for $p_1 - p_2$

Although non–informative or vague priors are often considered for a single proportion p, to our knowledge, no non–informative prior has been presented for the difference of two independent proportions.

Since there are at least two accepted non–informative priors for proportions, the uniform (or beta (1,1)) and Jeffrey's prior $\left(\text{or beta}\left(\frac{1}{2}, \frac{1}{2}\right)\right)$ various combinations of these two could give non–informative priors for p.

Case 1: beta $\left(\frac{1}{2}, \frac{1}{2}\right)$ – beta $\left(\frac{1}{2}, \frac{1}{2}\right)$ We then have, from the Theorem:

$f(p) = F_1\left(\frac{1}{2}, 0, \frac{1}{2}; 1; (1-p), 1-p^2\right)/\pi$ for $0 < p \leq 1$

and $f(p) = F_1\left(\frac{1}{2}, \frac{1}{2}, 0; 1; 1-p^2, 1+p\right)/\pi$ for $-1 \leq p < 0$

f has a vertical asymptote at $p = 0$ where it is not defined. This can also be seen from the above expressions of f since when either b_1 or b_2 equals zero $F_1(a, b_1, b_2; c; x, y)$ reduces to Gauss hypergeometric function $_2F_1(a, b; c; z)$, with $b = b_1$ or b_2 and it is known that $_2F_1$ converges when $z \to 1-$ only if $c - (a + b) > 0$.

Case 2: beta $\left(\frac{1}{2}, \frac{1}{2}\right)$ – beta(1,1). Then,

$f(p) = 2(p(1-p))^{\frac{1}{2}}\ F_1\left(\frac{1}{2}, 1, \frac{1}{2}; \frac{3}{2}; 1-p, 1-p^2\right)/\pi$ for $0 < p \leq 1$

$f(0) = 1$ and

$f(p) = 2(-p)^{\frac{1}{2}}(1+p)^{\frac{1}{2}}\ F_1(1, 0, 1; \frac{3}{2}; 1-p^2, 1+p)/\pi$ for $-1 \leq p < 0$

f is continuous at 0 and the same remark as in case 1 applies here to explain the behavior of f there.

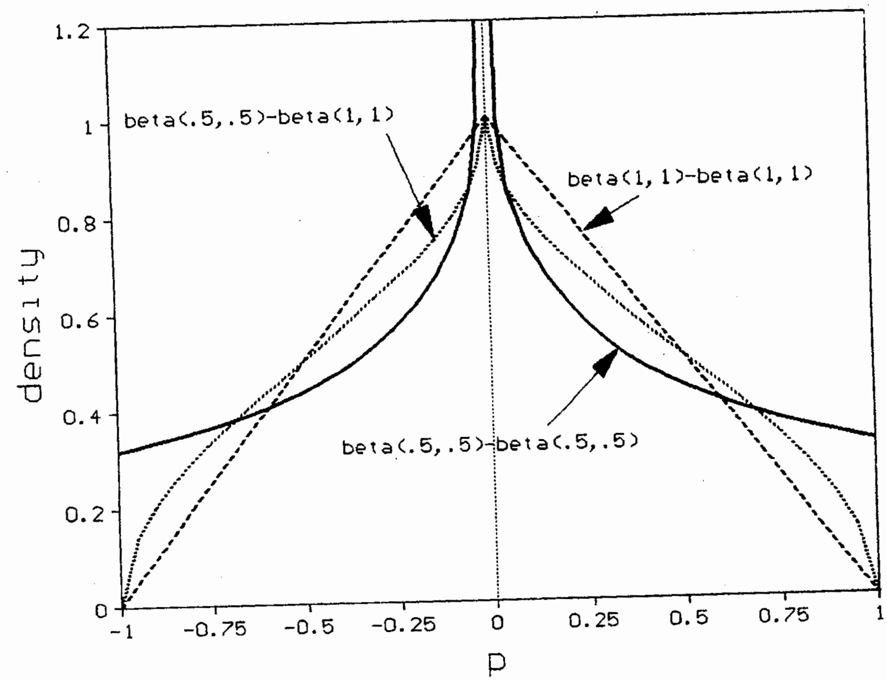

FIG. 1 Non–informative prior distributions for $p = p_1 - p_2$

<u>Case 3</u>: beta $(1, 1)$ – beta $\left(\frac{1}{2}, \frac{1}{2}\right)$

the graph here is identical to the graph of case 2.

<u>Case 4</u>: beta$(1, 1)$ – beta$(1, 1)$

Directly from equation (3) we can see that f is a triangular density with vertices at $(-1, 0)$, $(0, 1)$ and $(1, 0)$.

Fig. 1 gives the graph of f for different cases.

5. A numerical example

Let $p_1 \sim \text{beta}(3, 5)$ and $p_2 \sim \text{beta}(2, 8)$ be independent. The density of

$p = p_1 - p_2$, according to equations 2a, 2b and 2c, is then:

$$f(p) = 252\, p^{12} (1-p)^6\, F_1(5, 16, -2, 7, 1-p, 1-p^2) \text{ for } 0 < p \leq 1.$$

$$f(0) = \frac{18}{13}$$

and

$$f(p) = 21(-p)^{12} (1+p)^{10}\, F_1(8, -1, 16; 11; 1-p^2, 1+p) \text{ for } -1 \leq p < 0.$$

Let sampling results be $n_1 = 10$, $x_1 = 4$ and $n_2 = 6$, $x_2 = 2$. The posterior density of p_1 is then beta (7, 11) with mean $\mu_{\text{post}}(p_1) = \frac{7}{18}$ and variance $\text{Var}_{\text{post}}(p_1) = .012508$. For p_2 we have beta (4, 12), with $\mu_{\text{post}}(p_2) = \frac{1}{4}$ and $\text{Var}_{\text{post}}(p_2) = .011030$. The posterior density of p is:

For $0 < p \leq 1$, $f(p) = 185640\, p^{22}(1-p)^{14}\, F_1(11, 32, -6; 15; 1-p, 1-p^2)$ (9a)

$$f(0) = 1.67 \quad \text{and}$$

$$f(p) = 3337(-p)^{22} (1+p)^{18}\, F_1(12, -3, 32; 19; 1-p^2, 1+p) \text{ for}$$

$-1 \leq p < 0$ (9b)

The prior and posterior densities are given in fig. 2. We can now compute the 90% highest posterior density (hpd) credible interval for p using the computer program mentioned earlier and obtain $I_{.90} = (-.1124, .3922)$.

To study the robustness of the beta–difference distribution as a prior we introduce changes in the hyperparameters of the two beta priors. For the first prior several values between 2.5 and 3.5 are given to α_1 while β_1 takes various values between 4 and 6. Similarly, α_2 varies between 1.5 and 2.5

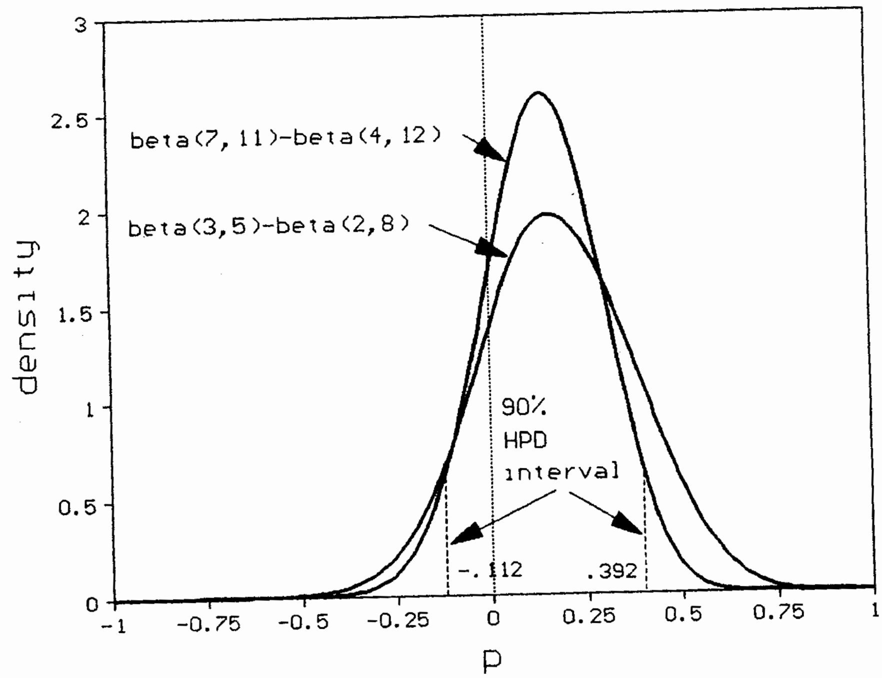

FIG. 2 Prior and Posterior distributions of $p = p_1 - p_2$

and β_2 between 4.5 and 7.5. Computations based on the corresponding equations (9a) and (9b) give credible intervals $I_{.90}$ which are just slightly different from the one above, with a difference in length not exceeding 10% of the latter's.

Without equations (9a) and (9b), using equation (5) with $c_{.95} =$ 1.645, the approximate 90% credible interval for p would be (–.1136, .3912), which is quite good an approximation.

The classical approach, based on sampling results only, gives $\hat{p}_1 = .40$, $\hat{p}_2 = .3333$, $\hat{\sigma}_1^2 = .0240$, $\hat{\sigma}_1^2 = .0240$ and hence a 90% confidence

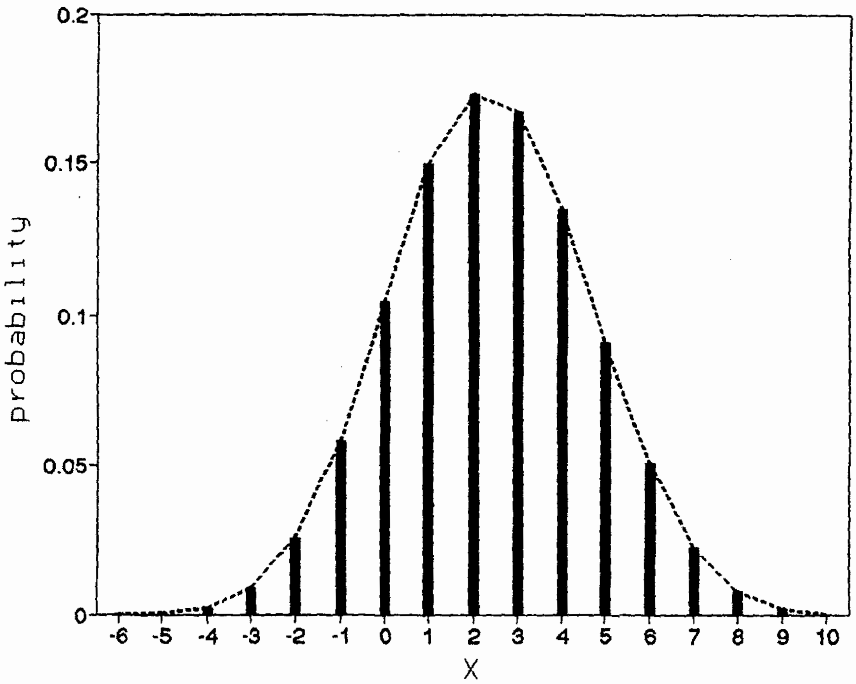

FIG. 3 Predictive distribution of $X = X_1 - X_2$

interval of (−.3398, .4730), which is considerably larger than either of the two credible intervals.

Finally, using equation (8) the predictive distribution of $X_1 - X_2$ is:

$$P(X = k) = \frac{1}{A} \sum_{l=m(k)}^{M(k)} \left[\begin{matrix}10\\ \ell\end{matrix}\right] \left[\begin{matrix}6\\ \ell-k\end{matrix}\right] B(3 + \ell, 15 - \ell)\, B(2 + \ell - k, 14 - \ell + k)$$

$k = -6, -5, \ldots, 10$, where $A = B(3,5)\, B(2,8) = 1/7560$, $m(k) = \max(0,k)$

Its graph is given by fig. 3.

ACKNOWLEDGMENT

This work was partially supported by NSERC Grant A9249 (Canada). The authors wish to thank Dr. Q. P. Duong of Bell Canada's Applied

Statistics Group (Montreal) for several helpful discussions and to an associate editor for several suggestions that have helped improving the content and presentation of this paper.

BIBLIOGRAPHY

Appell, P. and Kampé de Fériet, J. (1926). *Fonctions Hypergéométriques et Hypershériques. Polynomes d'Hermite*, Gauthier–Villars, Paris.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2^{nd} edition, Springer–Verlag, New York.

Draper, N. and Guttman, I. (1969). The value of Prior Information, in *New Developments in Survey Sampling*, N.L. Johnson and H. Smith ed., Wiley–Interscience, New York, 305–325.

Exton, H. (1976). *Multiple Hypergeometric Functions and Applications*, Chichester Ellis–Howard, London.

Irwin, J.O. (1927). On the Frequency–Distribution of the Mean of Samples from a Population Having any Law of Frequency with Finite Moments, *Biometrika* 19, 225–239.

Isaacs, G.L., Christ, D.E., Novick, M.R. and Jackson, P.M. (1974). *Tables for Bayesian Statisticians*, The University of Iowa Press, Iowa City.

Pham–Gia, T. and Turkkan, N. (1992). Information Loss in a Noisy Dichotomous Sampling Process, *Comm. Statist.-Theory Methods*, 21(7), 2001–2018.

Pham–Gia, T. (1990). A Bayes Sensitivity Analysis when using the Beta Distribution as a Prior (Comment on), *IEEE Trans. Reliability*, 39, 6–8.

Pham–Gia, T. and Duong, Q.P. (1989). The Generalized Beta and F–Distributions in Statistical Modelling, *Mathl. Comput. Modelling*, 13, 1613–1625.

Springer, M. (1979). *The Algebra of Random Variables*, John Wiley and Sons, New York.

Turkkan, N. and Pham–Gia, T. (1992). An algorithm to Compute HPD Intervals in Bayesian Analysis, *Journ. Stat. Comput. and Simul.*, 44, 243–250.

Wright, T. (1988). A note on Optimal Allocation and an Ordering of the Family of Beta Prior Distributions, *Comm. Stat. Theory–Meth.*, 17, 123–138.

Zacks, S. (1970). Bayesian Design of Single and Double Stratified Sampling for Estimating Proportion in Finite Populations; *Technometrics*, 12, 119–130.

Received December 1991; Revised January 1993