

Bayesian hypothesis testing for proportions

Antonio Nieto, PharmaMar, Madrid, Spain
Sonia Extremera, PharmaMar, Madrid, Spain
Javier Gómez, PharmaMar, Madrid, Spain

ABSTRACT

Most clinical trials contain tests on proportions, usually they are answered by means of the Frequentist approach, nevertheless another valid option could be to solve them using a Bayesian approach. The Bayesian approach has the advantage that it is not restricted to only one alternative hypothesis. Moreover, the hypotheses to be tested do not necessarily overlap. In this paper we show a SAS® macro to perform Bayesian hypothesis testing for proportions, that can be also extended to other kinds of endpoints and distributions. For simplicity only the null and one alternative hypothesis are shown. This macro is constructed assuming an improper prior distribution, the uniform (0,1), and a Beta as the posterior conjugate distribution. Therefore after calculating the proportion of successes in the trial, the probability of being under the null hypothesis or under the alternative hypothesis and a text label indicating the highest probability are shown.

INTRODUCTION

This paper has not the aim to confront Frequentist approach vs. Bayesian approach. In fact, both approaches can coexist and should be used indistinctly in the statistical interest. Consequently, we have implemented easy SAS macros to calculate the probabilities of different hypotheses using a Bayesian approach.

TESTS ON PROPORTIONS

Almost all, if not all, clinical trials contain tests on proportions. The proportion distribution is a collection of 'n' Bernoulli experiments; i.e., it is counted as the sum of the number of successes/failures out of 'n' independent samples.

Proportion tests usually are solved by means of a Frequentist approach, but this is not the only way. In a Frequentist analysis, if the comparison p-value is lower than the significance level selected, then the null hypothesis is rejected. In a Bayesian approach, the probability to be under any hypothesis is estimated and then these probabilities can be compared to decide what is the most plausible alternative.

THE BAYES' THEOREM

The Bayesian approach is based on the Bayes' theorem (1763), and expresses the conditional probability of a random event A given that an event B has occurred in terms of the conditional probability distribution of the event B given that A has occurred and the marginal probability of only A. In other words, beginning with the prior experience/knowledge (i.e., "a priori distribution") and then joining it with the trial investigation, a posterior conjugate distribution is obtained to be used to produce probabilities once the clinical trial has been completed.

BAYESIAN TESTS

The sum of the Bernoulli experiments is a Binomial distribution, which combined with the "a priori" information should lead to a posterior known distribution that allow an easy calculation of probabilities. Then, in practice, we will need to model the prior information to find the probability distribution that better fits the "a priori" knowledge and to lead to a posterior distribution easy to handle.

The Bayesian approach has the advantage that it is not restricted to only one alternative hypothesis. In addition, the hypotheses to be tested do not necessarily overlap and, therefore, probabilities associated under any hypotheses can be calculated in function of the different cutoffs selected as long as we know the conjugate distribution to be used.

PhUSE 2011

When the endpoint in a clinical trial follows a binomial distribution, the most appropriate distribution to model the “a priori” information is the Beta distribution. If we know that the prior probability of response can be modeled following a Beta distribution with parameters α and β , then it can be derived that the posterior conjugate Beta-Binomial distribution will have parameters $a=\sum x_i+\alpha$ and $b=n-\sum x_i+\beta$.

For the Bayesian test on proportions, the initial assumption is that the prior probability of the proportion could be any value between zero and one (i.e., no “a priori” information is available). In this case, an improper prior distribution, the uniform (0,1), can be assumed. From the uniform (0,1) or its equivalent beta (1,1) as prior distribution, the Beta-binomial distribution of the Bayes estimate under quadratic loss will follow a Beta distribution with parameters $a=\sum x_i+1$ and $b=n-\sum x_i+1$, being “ $\sum x_i$ ” the number of successes in the experiment and ‘n’ the number of independent samples (i.e., patients in a clinical trial).

The utility of the Bayesian tests is enhanced when some information is available on the parameter to be estimated before the clinical trial is started. The “a priori” assumption could be modified if the range of possible values where the proportion is contained, is previously known; in that case, we have to find the prior beta distribution that fits to the initial assumption and derive the Beta-Binomial conjugate distribution, as explained above.

For example, if a panel of experts concludes that the probability of response to a treatment in a clinical trial will fall between 0.3 and 0.7 with a high probability (i.e., 95%), and we want to perform a clinical trial to test hypotheses about the probability of response, then we need to find a Beta distribution that fits to this “a priori” information.

Since the beta distributions for values relatively high of α and β have an approximately normal shape, we can model easily a normal distribution and use the mean (m) and standard deviation (s) of the normal distribution to characterize the “a priori” beta distribution. In a normal distribution, we know that we can find >95% of the probability between $m-2s$ and $m+2s$. Therefore, we can make $m-2s=0.3$ and $m+2s=0.7$, and taking into account the symmetric shape of the normal distribution, $m=0.5$ and $s=0.1$. We just have to find a beta distribution with these mean and standard deviation, and one way to approximate could be by means of a moment's method type. As we know that for a beta distribution:

$$m=\alpha / (\alpha + \beta)$$
$$s^2=m(1-m) / (\alpha + \beta + 1)$$

and having the values of m and s, it can be derived that:

$$\alpha = [m^2 (1-m) / s^2] - m$$
$$\beta = (\alpha - m\alpha) / m = [m (1-m)^2 / s^2] + m - 1$$

In our example $M= 0.5$, $S= 0.1$, and we can calculate:

$$\alpha = [0.5^2 (1-0.5) / 0.1^2] - 0.5 = 12$$
$$\beta = [0.5 (1-0.5)^2 / 0.1^2] + 0.5 - 1 = 12$$

Then, the prior distribution is approximated as a Beta (12,12) and the conjugate binomial-beta distribution to be derived when we know the results from our clinical trial would be Beta($\sum x_i+12$, $n-\sum x_i+12$).

SAS MACRO (1)

Macro assuming a prior Uniform (0,1). In the macro after calculating the proportion of successes in the trial, the probability of being under the null hypothesis or under the alternative hypothesis and a text label indicating the highest probability are shown below.

```
/*
Bayesian macro to test two hypotheses with a non-informative prior distribution
(Uniform(0,1)=Beta (1,1));

-Variables needed
* x: number of successes in the sample
* n: sample size

H0: Null hypothesis
H1: Alternative hypothesis

-The conjugate distribution is a Beta(x+1,n-x+1)
*/
```

PhUSE 2011

```
%MACRO Bayes_test (x=,n=,H0=,H1=);

DATA bayes1;
length test $255.;
alfa=&x+1;
beta=&n-&x+1;
h0="p<="||left(trim(&H0));
h1="p>"||left(trim(&H1));
x=&x;
n=&n;

x1=probbeta(&H0,alfa,beta);
x2=1-probbeta(&H1,alfa,beta);

If x1>x2 then test='H0 is more probable than H1';
else if x1<x2 then test='H1 is more probable than H0';
else if x1=x2 then test='Equally probable hypotheses ';

run;

Proc print data=bayes1 noobs 1;
var h0 h1 x n test x1 x2;
label h0='H0' h1='H1' x='X' n='N' test='Test' x1='Prob. under H0' x2='Prob. under H1';
title "Bayes test of &x successes in &n samples";
footnote "Prior distribution Uniform (0,1)";
run;
title;
footnote;

%MEND Bayes_test;
```

EXAMPLE 1

We plan a clinical trial, with $n=40$ as sample size, no prior information on the proportion of responders, and we would like to test the hypotheses:

H_0 : Proportion of responders is $\leq 40\%$.

H_1 : Proportion of responders is $> 60\%$.

If we obtain 24 successes in our trial (i.e., 24 patients responding to a given experimental therapy), then we can obtain the posterior probability of the null and alternative hypotheses taking into account the results in our sample.

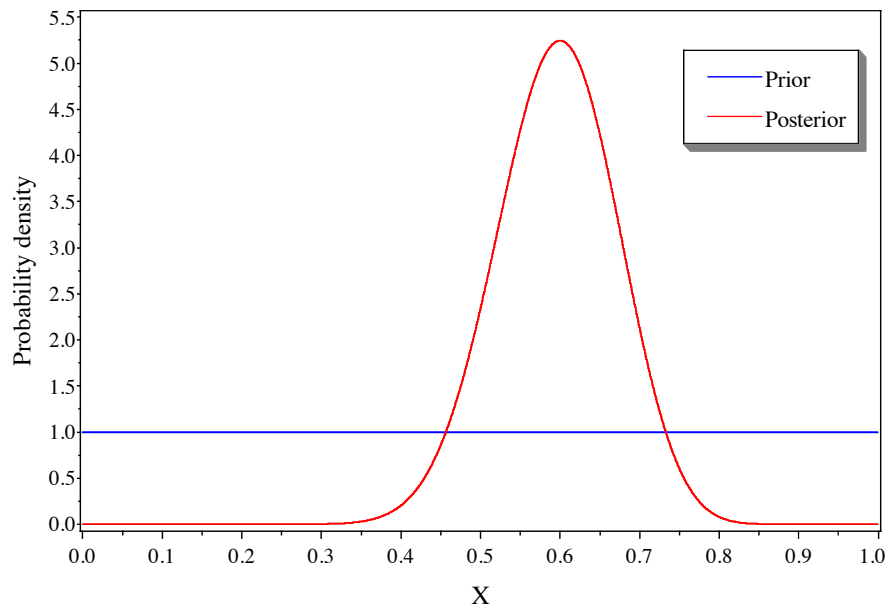
```
%Bayes_test (x=24,n=40,H0=0.40,H1=0.60);
```

Bayes test of 2 successes in 40 samples

H0	H1	X	N	Test	Prob. under H0	Prob. under H1
$p \leq 0.40$	$p > 0.60$	24	40	H1 is more probable than H0	0.005347226	0.48303

Prior distribution: Uniform (0,1)

Beta distribution plot



SAS MACRO (2)

Macro assuming a prior Beta (α , β). In the macro after calculating the proportion of successes in the trial, the probability of being under the null hypothesis or under the alternative hypothesis and a text label indicating the highest probability are shown below.

```
/*
Bayesian macro to test two hypotheses with a Beta prior distribution
(Beta (alpha,beta));

-Variables and parameters needed
* x: number of successes in the sample
* n: sample size
* Alpha: alpha parameter of the prior beta distribution
* Beta: beta parameter of the prior beta distribution

H0: Null hypothesis
H1: Alternative hypothesis

-The conjugate distribution is a Beta(x+alpha,n-x+beta)
*/

%MACRO Bayes_test (x=,n=,H0=,H1=,alpha=,beta=);

DATA bayes1;
length test $255.;
a=&x+&alpha;
b=&n-&x+&beta;
h0="p<="||left(trim(&H0));
h1="p>"||left(trim(&H1));
x=&x;
n=&n;

x1=probbeta(&H0,a,b);
x2=1-probbeta(&H1,a,b);

If x1>x2 then test='H0 is more probable than H1';
else if x1<x2 then test='H1 is more probable than H0';
else if x1=x2 then test='Equally probable hypotheses';

run;
```

PhUSE 2011

```
Proc print data=bytes1 noobs 1;
var h0 h1 x n test x1 x2;
label h0='H0' h1='H1' x='X' n='N' test='Test' x1='Prob. under H0' x2='Prob. under H1';
title "Bayes test of &x successes in &n samples";
footnote "Prior distribution Beta (&alpha,&beta)";
run;
title;
footnote;

%MEND Bayes_test;
```

EXAMPLE 2

In the same trial of Example 1, we know that the proportion of responders will fall within [0.3-0.7] with a 95% probability, and we would like to test the same hypotheses.

As calculated above, the prior distribution is Beta (12,12), and after obtaining 24 successes in our trial, thus the conjugate distribution is a Beta (24+12, 40-24+12) -> Beta (36, 28).

The macro call will be:

```
%Bayes_test (x=24,n=40,H0=0.40,H1=0.60,alpha=12,beta=12);
```

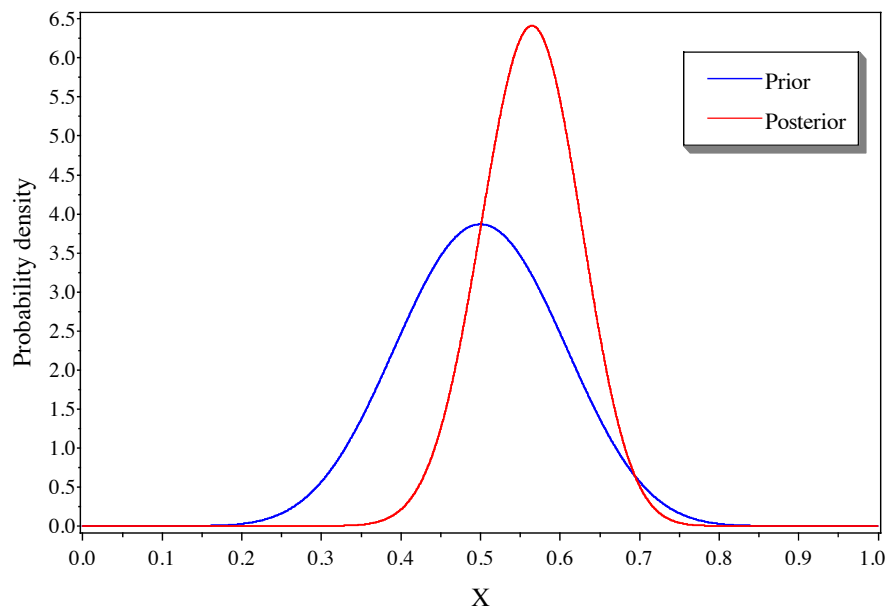
Bayes test of 2 successes in 40 samples

H0	H1	X	N	Test	Prob. under H0	Prob. under H1
p<=0.40	p>0.60	24	40	H1 is more probable than H0	0.004406341	0.27539

Prior distribution: Beta (12,12)

As we can see, in this second example, the probabilities under Ho and H1 changed according to the existing prior distribution.

Beta distribution plot



NOTE: If prior distribution selected in SAS Macro (2) is Beta (1,1), then the conjugate results found are the same than those obtained with the first macro; therefore, this macro could be generalized and used alone.

CONCLUSION

Frequentist approaches are usually employed in clinical investigation as they are a good method to conduct proportion tests, but they are not the unique method available. Bayesian tests, especially in the context of adaptive designs, are nowadays being increasingly used. We presented a Bayesian approach to be included in the statistical armamentarium to test proportion hypotheses.

In this paper, we show SAS macros to perform Bayesian hypothesis testing for proportions, but its use can be also extended to other endpoints and distributions. The most important aspects to take into account for Bayesian tests are a good selection of the distributions and a clear definition of the "a priori" information collected.

REFERENCES

-SAS Online Doc.

-Bayesian Approaches to clinical trials and health care evaluation. David J. Spiegelhalter, Keith R. Abrams and Jonathan P. Myles. John Wiley and Sons. Dec 1, 2003.

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Antonio Nieto Archilla
Clinical Development. PharmaMar S.A.
Avda. de los Reyes, 1
Polígono Industrial La Mina
28770 Colmenar Viejo. Madrid (SPAIN)
anieto@pharmamar.com

Sonia Extremera Tenaguillo
Clinical Development. PharmaMar S.A.
Avda. de los Reyes, 1
Polígono Industrial La Mina
28770 Colmenar Viejo. Madrid (SPAIN)
sextremera@pharmamar.com

Javier Gómez García
Clinical Development. PharmaMar S.A.
Avda. de los Reyes, 1
Polígono Industrial de la Mina
28770 Colmenar Viejo. Madrid (SPAIN)
jgomez@pharmamar.com