Available online at www.sciencedirect.com
SCIENCE DIRECT®
Meat Science 69 (2005) 115–122
MEAT SCIENCE
www.elsevier.com/locate/meatsci

The use of Bayesian statistics in meat quality analyses: a review

A. Blasco*

Departamento de Ciencia Animal, Universidad Politécnica de Valencia, P.O. Box 22012, 46071 Valencia, Spain

Received 2 February 2004; received in revised form 15 June 2004; accepted 15 June 2004

Abstract

A Bayesian approach to meat quality analyses is discussed and compared with the classical statistical approach. Inferences from means, medians and modes of marginal posterior distributions are presented, and a variety of probability inferences and confidence intervals are presented and discussed. Classical and Bayesian theories of hypothesis testing are compared and their advantages and disadvantages are discussed. Fundamentals of Bayesian inference and Monte-Carlo Markov Chain (MCMC) techniques are presented and discussed. The great flexibility for inferences introduced by MCMC techniques is stressed. Practical examples of meat quality analyses are given, with references to available free software to analyze a large variety of models.

Keywords: Bayesian inference; MCMC; Meat quality

1. Introduction

Meat quality analyses are usually performed with small samples, mainly because it is expensive to measure a large amount of variables, some of them involving a lot of lab work or involving a trained panel test. Until now, classical statistics has been the usual way of expressing uncertainty in meat quality analysis whereas Bayesian analyses have been mainly applied by animal breeders to complicated genetic problems (see Blasco, 2001, for a review). However, Bayesian techniques have some advantages in comparison with classical methods. Bayesian inference is based in probabilities, which gives a great flexibility to construct all kinds of confidence intervals. Besides, Bayesian methods can concentrate the inference in the parameter of interest, by integrating out all other unknowns as we will see later. The generalised use of Bayesian statistics in many fields of science is relatively new. Although Bayesian methods were theoretically powerful, they usually led to formulas in which multiple integrals had to be solved in order to perform Bayesian inference. As these integrals could not be calculated, even using approximate methods, Bayesian inference had a limited range of applications. When Monte-Carlo Markov Chain (MCMC) methods appeared in the nineties and were applied to estimate marginal posterior distributions, the computation problems were solved and interest in Bayesian methods was renewed. Only recently there has been developed friendly software in MS-Windows to perform Bayesian analysis using MCMC techniques (Spiegelhalter, Thomas, & Best, 2000). Nowadays the use of MCMC techniques facilitate all the inferences needed from posterior distribution and their popularity has enormously increased in many fields of science. The objective of this paper is to explain the possibilities of Bayesian inference for meat quality analyses.

2. Describing uncertainty

Uncertainty is typically described in classical inference by giving a standard error (s.e.) of the estimator.

* Tel.: +34-963877433; fax: +34-963877439.
E-mail address: ablasco@dca.upv.es.

0309-1740/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.meatsci.2004.06.012

When the sampling distribution of the estimator is normal, approximately twice the s.e. is a confidence interval at 95% of probability. Notice that when a confidence interval is given, for example [5.60, 5.70] for a pH, this does not mean that the probability of the pH of being between 5.60 and 5.70 is 95%. What we say, is that if the experiment would be repeated an infinite number of times, we would get an infinite number of confidence intervals (of which [5.60, 5.70] is just an example) that would contain the true value of the pH in 95% of cases. As we only have one confidence interval from our experiment, we say that this interval is one of the intervals that contains the true value of the pH, with the hope that behaving like this in all experiments, we will be wrong only 5% of the times. Neyman and Pearson (1933) suggested that the 'scientific behaviour' should be to act as if the interval obtained was the true one, being sure that 'in the long run' we will be right in 95% of cases.

In a Bayesian context, the objective is to describe the uncertainty about the true value of some parameter using probability as a measure of this uncertainty. For example, if the parameter of interest is the difference between two treatments S and C for a meat quality trait, the aim of Bayesian inference is to find a probability density of the difference between treatments given the data, $P(S-C|\mathbf{y})$, where $\mathbf{y}$ is the vector of observations (Fig. 1). If we have many data, the probability will be concentrated around the true value and this distribution will be sharp, otherwise it will be flatter representing the uncertainty, given our data, that we have about the difference between treatments.

Once this distribution is obtained, inferences can be made in multiple manners. We can give several values of $S-C$ calculated from the distribution $P(S-C|\mathbf{y})$. We can give the mode, which is the maximum of the distribution; i.e., the most probable value of $S-C$ that we can infer with our data. We can also give the median, which has the property that the probability of the true value of $S-C$ being greater than the median is the same as being smaller. We can also give the mean of the distribution, which has the property of minimizing the risk of the estimation when the loss function is a quadratic one, which is the most common loss function used in statistics (see Section 6). Posterior distributions tend to be normal when the number of data increases. When the distribution is symmetrical, the three values are coincident.

Bayesian inference provides probability intervals. Now, the confidence intervals (Bayesians prefer to call them credibility intervals) contain the true value with a probability of 95%, or other probabilities defined by the user. An advantage of the Bayesian approach through MCMC procedures is the possibility of easy construction of all kind of intervals. This allows us to ask questions that we could not ask within the classical inference approach. For example, if we give the median and the mode and ask for the precision of our estimation, we can find the shortest interval with 95% probability of containing the true value (what is called the *Highest posterior density interval* at 95%), but here this interval is not dependent on the estimate we give, and it can be asymmetric about the mean or the mode (Fig. 1(a)). Of course, in the Bayesian case we can also obtain the symmetric interval about the mean or the mode containing 95% of the probability (Fig. 1(b)).

We can also calculate the probability of the difference between S and C being higher than 0 (Fig. 2(a)); i.e., the

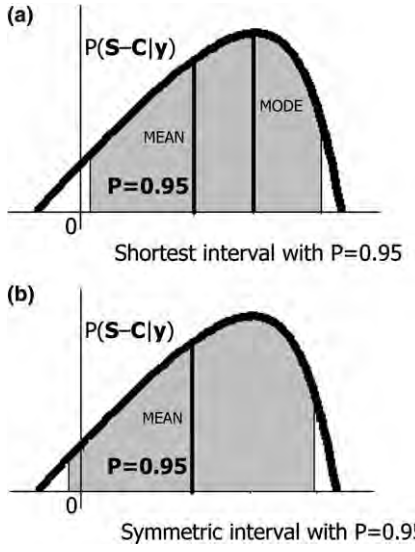

Fig. 1. Examples of credibility intervals of the difference between two treatments S and C containing the true value with a probability of 95%.

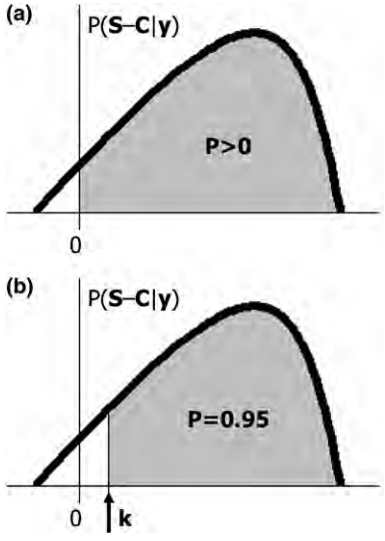

Fig. 2. Probabilities of events. (a) Probability of group S being higher than group C, (b) probability of groups S and C differing by at least a quantity k.

probability of S being greater than C (actually this is equivalent to a test of hypothesis). In the case in which S is less than C we can calculate the probability of $S-C$ being negative; i.e., the probability of S being less than C. In some cases it may be important to know how big we can state that this difference is with a probability of 95%. By calculating the interval $[k,+\infty)$ containing 95% of the probability (Fig. 2(b)) we can state that the probability of $S-C$ being less than this value k is only 5%; i.e., we can state that $S-C$ takes *at least* a value k with a probability of 95% (or the probability we decide to take). If S is lower than C, we can calculate the interval $(-\infty, k]$ and state that the probability of $S-C$ being higher than k is only 5%.

In practice, we are interested not only in finding whether S is higher than C or not, but whether this difference is relevant. Here 'relevant' means economic or biological relevance. S may be higher than C, but this difference may be irrelevant. We can calculate the probability of this difference being relevant. For example, if we are measuring lean content in pigs, we can consider that 1 mm of backfat is a relevant difference between S and C groups, and calculate the probability of $S-C$ being more than 1 mm (Fig. 3(a)). For many traits, it is difficult to state which is a relevant difference. For example, if we measure the effect of selection on enzymes activities it is not clear what we can consider to be 'relevant'. For these cases, we can express our results as ratios instead of differences. For example, we can calculate the probability of the selected population being 10% higher than the control population for a particular trait. We will do the same as before, but now using the marginal posterior distribution of the ratio S/C instead of the distribution of $S-C$ (Fig. 3(b)). We can also calculate the same credibility intervals as before for the posterior distribution $P(S/C|\mathbf{y})$.

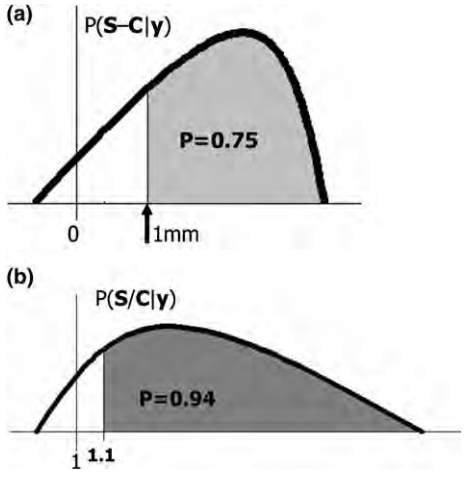

Fig. 3. Probability of relevance. (a) Probability of a difference in backfat between groups S and C being higher than a relevant value of 1 mm. (b) Probability of the enzyme activity of group S being 10% higher than the activity in group C.

3. Testing hypothesis

Having the posterior distribution of the difference between treatments, it is not necessary to perform a hypothesis test, because the question "is treatment S different from treatment C?" is answered by finding the probability of $S-C$ being higher than zero, given the data (or the probability of S/C being higher than 1, given the data). However, when correcting for some nuisance parameter (for example, the number of different sessions in a panel test) we can have many levels, and it is easier to perform a hypothesis test to determine whether this nuisance parameter is relevant or not. In classical statistics, assuming a risk of rejecting the hypothesis when it is true, regions of acceptance or rejection can be defined, based on the distribution of the sample under the hypothesis to be tested when the experiment is repeated an infinite number of times. The hypothesis is accepted or rejected depending on whether the actual sample falls in one or in the other region. Once the hypothesis is accepted, the 'scientific behaviour' should be acting as if this hypothesis were true. Note that the method gives no clues about how probable is one hypothesis related to the other. The answer given by a test of hypothesis is just 'yes' or 'no', because the risk and the corresponding regions are fixed *before* the experiment is performed. This produces some problems associated with common practice. For example, a hypothesis may be that two breeds have the same mean for some trait. Note that by augmenting sample size enough, this hypothesis will always be rejected, thus in a well-designed experiment, it should be decided when a hypothesis should be rejected; i.e., how large the difference between means should be, and we should use a sample size that will detect significant differences when they are relevant, and lead to non significant differences otherwise. However, as experiments are designed to find significant differences for one trait but many traits are usually measured, the problem arises in tests for traits not considered when designing the experiment. In all these cases, significance will not be associated to relevant differences because these differences were not considered when the experiment was designed. Thus, the answer provided by the test (to accept or reject the hypothesis) might be unsatisfactory, since we can find irrelevant differences that are significant, or relevant differences that are not significant. Moreover, there are traits for which it is difficult to establish what a "relevant difference" can be, for example for enzyme activities or for panel test results. Another common practice problem is the misinterpretation of P-values. Sometimes P-values (the probability of the present results when the null hypothesis holds) are wrongly interpreted as significance levels, or as the probability of the null hypothesis being true. However, the risk of the hypothesis is fixed *before* the experiment is performed, and if we repeat the

experiment we know we will find another *P*-value different from the *P*-value we have obtained. Thus we cannot consider the *P*-value as a significance level because in classical statistics we make inferences not only from our sample but from what would happen after infinite repetitions of the experiment.

The result of a test of hypothesis is different in the Bayesian School. In a Bayesian framework the probability of each hypothesis is calculated and compared. For example, if the hypothesis to be tested (H_0) is that two breeds have the same mean for a trait, and the alternative (H_1) is that the means are different, in a Bayesian context we can say that H_0 is seven times more probable than H_1. Not only the null and the alternative hypotheses, more hypotheses can be considered. When a Bayesian procedure is performed, instead of prior risks we have the actual probabilities of each of the hypothesis given the data, and we can quantify how much evidence the data give in favour of a hypothesis. However, there are some difficulties in applying these tests in practice, and some alternatives have been suggested, as we will see later.

4. Bayesian inference and MCMC

As we said before, Bayesian inference is based in probability densities of the unknowns that are estimated, given the data. To find these probability densities, we will use the laws of probability. The probability of two events happening together is

$$P(\mathrm{A},\mathrm{B}) = P(\mathrm{A}|\mathrm{B}) \cdot P(\mathrm{B}) = P(\mathrm{B}|\mathrm{A}) \cdot P(\mathrm{A}),$$

where $P(\mathrm{A}|\mathrm{B})$ means the probability of A given that B occurred. From this expression, we obtain

$$P(\mathrm{A}|\mathrm{B}) = P(\mathrm{B}|\mathrm{A}) \cdot P(\mathrm{A})/P(\mathrm{B}).$$

This later expression is known as Bayes theorem. If we applied Bayes theorem to the probability of the difference between treatments *S* and *C* given the data **y**, we obtain

$$P(S-C|\mathbf{y}) = P(\mathbf{y}|S-C) \cdot P(S-C)/P(\mathbf{y}),$$

where $P(\mathbf{y})$ is a constant (the probability of the sample), $P(\mathbf{y}|S-C)$ is the distribution of the data given the unknowns (usually taken as normal), and $P(S-C)$ is the probability of the unknowns that does not depend on our data. The criticism of the Bayesian inference is focused on this last probability, called probability a priori because it should be defined independently of the data and before the start of the experiment. Sometimes this prior probability is well defined (e.g., in genetics, the *prior probability* of obtaining a recessive individual when crossing two heterozygous is 1/4 according to the Mendel's law), but more often this probability is not clearly established. Bayesian statisticians tend to use $P(S-C)$=*constant* (these prior distributions are called *flat priors*), or other neutral mathematical functions to avoid the problem. Then, $P(\mathbf{y}|S-C)$ dominates Bayes theorem and the inferences are essentially based on the data. A detailed account of Bayesian procedures and MCMC can be found in Sorensen & Gianola (2002).

To find these *marginal* posterior distributions and to calculate all kind of credibility intervals that we have mentioned before, multiple integrals had to be solved that often could not be calculated, even using approximate methods. A powerful technique known as MCMC was developed in the nineties in order to make possible Bayesian inference from marginal posterior distributions. When using MCMC, random samples of the marginal posterior distributions are obtained. The result of a MCMC analysis are lists of numbers (called *chains*) randomly extracted from the marginal posterior distribution of interest. For example, we can obtain a *chain* for treatment *S* and another *chain* for treatment *C*, for intensity of rabbit odour evaluated by a panel test:

S: [3.1, 3.3, 4.1, 4.8, 4.9,...]
C: [2.4, 2.6, 2.6, 2.6, 2.8,...]

these are samples of the marginal posterior distributions of *S* and *C* treatments respectively. From them we can calculate the *chain* for the Marginal posterior distribution of $S-C$ or of S/C by subtracting or dividing the samples of the chains.

$S-C$: [0.7, 0.7, 1.5, 2.2, 2.1, ...]
S/C: [1.3, 1.3, 1.6, 1.8, 1.8, ...]

With these chains we can construct histograms as fine as we want (Fig. 4), because we can take as many samples of the distribution as we desire, and make the inferences shown in Figs. 1–4. For example, to estimate the mean of the distribution, we just take the average of

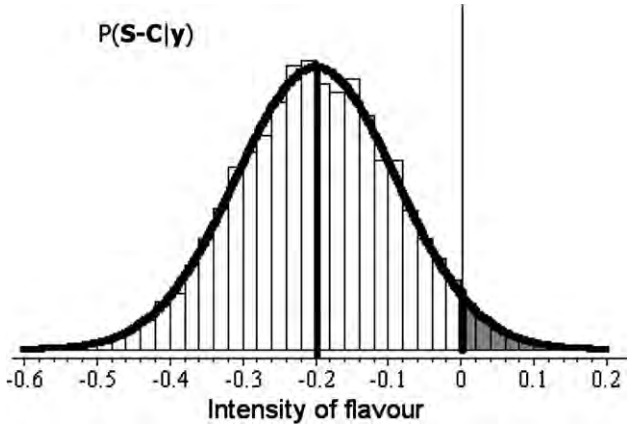

Fig. 4. Histogram from a MCMC chain estimating the marginal posterior distribution, with features of Bayesian inference.

the chain. To estimate the median we sort the chain and take the sample that is in the middle. To estimate the mode, we draw the density function and take the maximum. To estimate probabilities we sort the chain and calculate the proportion of points over some sample; for example, to calculate $P(S-C|\mathbf{y}) > 0$ we calculate the proportion of positive points of the chain for $S-C$. To calculate intervals we find the samples delimiting 95% of the points of the chain; for example, to calculate the interval $[k, +\infty)$ we order the chain from the highest to the lowest value and take the 95% of the points with the highest value, k being the last sample. To calculate the shortest interval containing 95% probability (HPD 95%), we order the chain as before and take all pairs of samples containing 95% of the points, selecting the couple of samples that are closest.

As we are not using the true marginal posterior density, but an estimation of it, we have a sampling error that depends on the number of samples of the chain. This error (called *Monte-Carlo error*) can be made as small as we desire because we can take as many samples as we want. The common practice is to take many samples (say, 10,000 for the common analyses) and give the value of the *Monte-Carlo standard error*, which should then be very small. MCMC procedures are iterative, thus another feature of the chain that is usually given is the result of some *convergence test*, which ensures we are taking samples of the right posterior density function (an intuitive description of the procedure can be found in Blasco, 2001). Due to the recent development of MCMC techniques, only recently has such software in MS-Windows been available. The Medical Research Council and the Imperial College of Science, Technology and Medicine have developed free friendly software for Bayesian analyses using MCMC techniques (Spiegelhalter et al., 2000, http://www.mrc-bsu.cam.ac.uk/bugs), performing all the features of the marginal posterior distributions needed for a wide range of models.

5. Bayesian analysis in practice

Bayesian inference modifies the approach to the discussion of the results. Classically, we have point estimation, usually a least square mean, and its standard error, accompanied by a hypothesis test indicating whether there are differences between treatments. Then we discuss the results based upon these features. Now the procedure is inverted and we ask first which question is relevant for us and go to the marginal posterior distribution to find an answer.

Example 1. We take an example by Blasco et al. (1994). They were interested in finding the differences in percentage of ham of a pig cross using Belgian Landrace or Duroc as terminal sire. They offer least square means of $25.1 \pm 0.2\%$ and $24.5 \pm 0.2\%$ respectively and find that they are significantly different. Now, in order to present the Bayesian results we have estimated the marginal posterior distribution of the difference between both crosses. Now, we should ask some questions:

(1) *What is the difference between both crosses?*

We can offer the mean, the mode or the median. Here the marginal distribution is approximately Normal, thus the three parameters are the same, and the answer is coincident with the classical analysis: 0.6%.

(2) *What is the precision of this estimation?*

The most common Bayesian answer is the Highest posterior density interval containing a probability of 95%. But here the marginal posterior distribution is approximately normal, thus we know that ±twice the standard deviation of the marginal posterior distribution will contain approximately 95% of the probability, thus we can either give an interval [0.1%, 0.11%] or just the standard deviation of the difference, 0.25%.

(3) *What is the probability of the Belgian Landrace cross being higher than the Duroc cross?*

We do not need a test of hypothesis to answer this question. We can just calculate how much probability area is positive of the marginal posterior distribution. We calculate the proportion of points of the Gibbs sampling chain that is positive. We find 99% of probability. Note that we could have found a high posterior density interval containing 95% of probability of [0.0%, 0.12%], if for example the standard deviation would have been 0.30%, and still say that the probability of the Belgian landrace cross being higher than the Duroc is 97%.

(4) *How large can we say is this difference with a probability of 95%?*

We calculate the interval $[k, +\infty)$ by ordering the Gibbs sampling chain as before and finding the value in which 95% of the points are contained. As our chain has 10,000 samples, we go to the sample number 500 and find that the value of this sample is 0.2%, thus we can say that the difference between crosses is at least 0.2% with a probability of 95%.

(5) *Considering that an economically relevant difference between crosses is 0.5%, what is the probability of the difference between crosses being relevant?*

We calculate the proportion of samples of the Gibbs sampling chain that are higher than 0.5, and we find the value to be 66%. Thus, we can say that although both crosses are different, the probability of this difference being relevant is 66%.

Example 2. We now take a sensory analysis from Hernández et al. (2004). Here a population selected for growth rate is compared with a control population, and sensory properties of meat from the *l. dorsi* are assessed by a test panel. The panel scored from 0 to 10, and data were divided by the standard deviation of each panelist in order to avoid a scale effect. In this example, it is

difficult to determine what a relevant difference is, thus instead of assessing the differences between the selected (S) and control (C) population, the ratio of the selection and control effects S/C is analyzed. This allows us to express the superiority of the selected over the control population (or conversely the superiority of the control over the selected population) as a percentage. We will take the trait liver flavor. The result of the classical analysis is that the least square means of the selected and control populations are 1.38 ± 0.08 and 1.13 ± 0.08 and they were found to be significantly different. These means and their standard error are rather inexpressive about the effect of selection on meat quality. Now, the Bayesian analysis answers the following questions:

(1) *What is the probability of the selected population being higher than the control population?*

We calculate the proportion of points of the Gibbs sampling chain that are higher than 1. We find 99% probability.

(2) *How much higher is the liver flavor of the selected population with respect to the control population?*

As in Example 1, we can give the mean, the mode or the median, and as the marginal distribution is also approximately normal all of them are coincident, we find that the liver flavor of the selected population is 23% higher than the liver flavor of the control population.

(3) *Which is the precision of this estimation?*

The 95% high posterior density interval goes from 1.03 to 1.44, which means that the liver flavor of the selected population is between 3% and 44% higher than this flavor in the control population with a probability of 95%.

(4) *How large can we say is this difference with a probability of 95%?*

We calculate the interval $[k, +\infty)$ by ordering the Gibbs sampling chain as before and finding the value in which 95% of the points are contained. The chain of Hernández et al. (2004) had 8000 samples, thus we go to sample number 400 and find that the value of this sample for the ratio S/C is 1.06, thus we can say that selected population is at least 6% higher than the control population with a probability of 95%, or that the probability of selected population being lower than 6% of the control population has a probability of only 5%.

(5) *Considering 10% higher as being relevant, what is the probability of the selected population being 10% higher than the control population?*

We calculate the proportion of samples of the Gibbs sampling chain that are higher than 1.10 and find this value to be 88%. This means that the probability of the effect of selection on liver flavor being relevant is 88%. This is not related to significance thresholds or rejection areas, we can state that this is the actual probability, and it is a matter of opinion whether this probability is high or low.

Example 3. Hernández, Pla, and Blasco (1998) estimated the correlation between moisture and fat percentage in the hind leg of rabbit, obtaining a coefficient of -0.95 ± 0.07. This standard error is not very useful, since the sampling distribution of the correlation coefficient is not symmetrical (the correlation cannot be lower than -1, thus the $\pm$ is misleading). A Bayesian analysis obtained the marginal posterior distribution shown in Fig. 5. Here the distribution is asymmetrical, thus mode, mean and median are not coincident. A usual choice is to take the mean (-0.93) because it minimizes the quadratic risk, which is conventional, although there are reasons to take the median or the mode (see Section 6). Here the HPD interval at 95% is $[-1.00, -0.79]$, not symmetrical around the mean, and shows better the uncertainty about the correlation than the s.e. of the classical analysis. The probability of this correlation being negative is one, as expected.

Another case in which asymmetric distributions are found is when estimating residual standard deviations. A usual procedure in classical statistics, rather inconsistent, is to use an unbiased estimator for the variance and then to calculate its square root to find the standard deviation. The problem is that the square root of an unbiased estimator does not give an unbiased estimator any more, and then the standard deviation obtained is a *biased* estimator of the standard deviation. In Bayesian statistics there is no bias because there are not infinite repetitions of the experiment. MCMC procedures guarantee that obtaining a chain for the marginal posterior distribution of the variance and calculating the square root of each sample of the chain we will obtain a chain for the standard deviation that will estimate its marginal posterior distribution. If we take the median of this instead of the mean, the square of this median will be the median of the marginal posterior distribution for the variance, which is an interesting property (see Section 6). If we are not interested in the variance but only

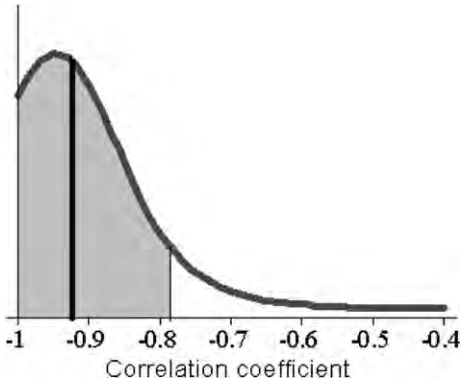

Fig. 5. Marginal posterior distribution of the correlation between moisture and fat percentage of the hind leg of rabbits.

in the standard deviation, we can take the mean of the marginal posterior distribution, the median or the mode, depending on our preferences.

6. Discussion

Bayesian techniques have been criticized because they need to integrate prior information, and this is difficult, in general. Because of this, modern Bayesians tend to consider the prior distribution just as a mathematical artefact that allows them to work with probabilities (see, for example, Bernardo & Smith, 1994). Bernardo has proposed the use of priors that produce posterior distributions with minimal prior information (called *reference priors*, see Bernardo & Smith, 1994). Although these priors are usually difficult to calculate, for models normally used in meat quality analyses, reference priors are well defined and easy to implement. Even if these reference priors are not used, bounded flat priors within limits can be used to give the same results in practice.

When posterior distributions are asymmetrical, mean, median are mode will give different estimators of the unknown of interest, and it is a matter of preference to take one or the other. Means minimize the quadratic risk, which rather conventionally is accepted as the right risk to minimize. Medians minimize the risk when the loss function is the absolute value of the difference between the true value and the estimate rather than its square, and has the property of being invariant to one-to-one transformations, which is useful for example when a coefficient of correlation and its square are used in the same work. Modes have the attractive property of being the most probable value, but it minimizes the risk when the loss function is one if the estimate is the true value and zero otherwise, which is a poor loss function.

An advantage of Bayesian inference is that by using probabilities, inferences are made from the *marginal* posterior distribution, having integrated out, weighing by their probabilities, all the other unknown parameters. For example, in our example about treatments S and C, we do not know the residual variance, that has to be estimated from the data. Suppose that this residual variance can only take two values, $\sigma^2 = 0.5$ and $\sigma^2 = 1$. The marginal posterior distribution of the difference between treatments will be the sum of $P(S - C$ given the data and given that $\sigma^2 = 0.5)$ multiplied by the probability of σ^2 taking the value 0.5, plus $P(S - C$ given the data and given that $\sigma^2 = 1)$ multiplied by the probability of σ^2 taking the value 1.

$$P(S - C|\mathbf{y}) = P(S - C|\mathbf{y}, \sigma^2 = 0.5)P(\sigma^2 = 0.5) + P(S - C|\sigma^2 = 1)P(\sigma^2 = 1).$$

When σ^2 can take all possible values from 0 to ∞, instead of summing up we calculate the integral of $P(S - C|\mathbf{y}, \sigma^2)$ for all σ^2 from 0 to ∞. Thus, we take all possible values of the unknowns, we multiply by their probability and we sum up, which means we take into account all possible values that the other unknowns can take and we concentrate our efforts of estimation only in the posterior probability of the unknown of interest.

Hypothesis test is an open area because some inconsistencies have been discovered in both classical and Bayesian schools. This has led to a new area of inference called "model choice", in which the problem to be solved is to select a model among several choices (including or not an effect, for example, or assuming a distribution or another for the data). Several criterions with different properties have been proposed for model choice. Recently, Spiegelhalter, Best, Carlin, and Linde (2002) have proposed a new criterion within the Bayesian frame, with a wider scope, that can be easily calculated from MCMC results. This criterion, called deviance information criterion (DIC) selects the model with the highest chance of predicting a replicate data set, and it is given in the standard software.

Care should be taken with MCMC techniques when using uncommon distributions, or in complicated multivariate models, but in these last cases, classical analyses should also be carefully performed as problems with them may also appear. No problems should appear for most meat quality analyses, of the type we have dealt with in this paper.

Bayesian analysis is a powerful tool for inference, with great flexibility to present uncertainty about the results of meat quality experiments. Nowadays there is friendly software for easily performing a large part of the analyses required by meat scientists, and it can be predicted a quick development in meat science, similar to the effects of these techniques have had in other fields of biology.

Acknowledgements

I am grateful to Prof. J.M. Bernardo and Dr. L. Varona for useful comments and advise. This work was supported by MCYT Project No. AGL2002-04383-C02.

References

Bernardo, J. M., & Smith, F. M. (1994). *Bayesian theory*. New York: Wiley.

Blasco, A. (2001). The Bayesian controversy in Animal Breeding. *Journal of Animal Science*, 79, 2023–2046.

Blasco, A., Gou, P., Gispert, M., Estany, J., Soler, Q., Diestre, A., et al. (1994). Comparison of five types of pig crosses. I. Growth and Carcass traits. *Livestock Production Science*, 40, 171–178.

Hernández, P., Guerrero, L., Ramírez, J., Mekkawy, W., Pla, M., Ariño, B., et al. (2004) A Bayesian approach of the effect of

selection for growth rate on sensory meat quality of rabbit. Proceedings of the Eighth World Rabbit Congress, Mexico, 7–10 September 2004.

Hernández, P., Pla, M., & Blasco, A. (1998). Carcass characteristics and meat quality of rabbit lines selected for different objectives. II. Relationships between meat characteristics. *Livestock Production Science*, 54, 125–131.

Neyman, J., & Pearson, E. (1933). On the problem of the most efficient test of statistical hypotheses. *Philosophical Transactions of the Royal Society A*, 231, 289–337.

Sorensen, D., & Gianola, D. (2002). *Likelihood, Bayesian, and MCMC methods in quantitative genetics*. New York: Springer.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. van der (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society B*, 64, 583–639.

Spiegelhalter, D. J., Thomas, A., & Best, N. G. (2000). WinBUGS Version 1.3 User Manual. Cambridge: Medical Research Council Biostatistics Unit. Available from http://www.mrc-bsu.cam.ac.uk/bugs.