

Editorial Manager(tm) for Nursing Research
Manuscript Draft

Manuscript Number: NRES-D-07-00107

Title: Bayesian Data Analysis: Estimating the Efficacy of Tai Chi as a Case Study

Article Type: Method

Corresponding Author: Dr. Byron Gajewski, PhD

Corresponding Author's Institution: University of Kansas Medical Center

First Author: Jacque Carpenter , RN, PhD

Order of Authors: Jacque Carpenter , RN, PhD; Byron Gajewski, PhD; Cynthia Teel, RN, PhD; Lauren S Aaronson, RN, PhD, FAAN

Manuscript Region of Origin: UNITED STATES

Abstract: Background: Bayesian inference provides a formal framework for updating knowledge by combining prior knowledge with current data. Over the past ten years, the Bayesian paradigm has become a popular analytic tool in health research and this popularity is likely to increase in the near future. While the nursing literature contains examples of Bayes' theorem applications to clinical decision-making, it lacks an adequate introduction to Bayesian data analysis.

Approach: Bayesian data analysis is introduced through a fully Bayesian model for determining the efficacy of T'ai Chi as an illustrative example. The mechanics of using Bayesian models to combine prior knowledge, or data from previous studies, with observed data from one's current study are discussed.

Results: The primary outcome in the illustrative example was physical function. Three prior probability distributions (priors) were generated for physical function using data from a similar, previous study found in the literature. Each prior was combined with the likelihood from observed data in the current study to obtain a posterior probability distribution. In each case, the posterior distribution showed that the probability that the control group is better than the T'ai Chi treatment group was low.

Discussion: Bayesian analysis is a valid technique that allows the researcher to manage varying amounts of data appropriately. As advancements in computer software continue, Bayesian techniques will become more

accessible. Researchers must educate themselves on applications for Bayesian inference, as well as its methods and implications for future research.

Dear Dr. Dougherty,

On behalf of my co-authors, please accept our submission “Bayesian Data Analysis: Estimating the Efficacy of Tai Chi as a Case Study” for consideration to be published in *Nursing Research Methodology*. We sent you a query regarding the potential for the idea of this paper and you encouraged us to submit.

This paper is not being considered for publication in any other journal.

Sincerely,

Byron Gajewski

Running Head: BAYESIAN DATA ANALYSIS: ESTIMATING THE EFFICACY OF
T'AI CHI

Bayesian Data Analysis: Estimating the Efficacy of Tai Chi as a Case Study

Jacque Carpenter, RN, PhD
Clinical Assistant Professor
University of Kansas School of Nursing
3901 Rainbow Blvd.
Kansas City, Ks 66160

Byron Gajewski, PhD
Assistant Professor
University of Kansas Schools of Nursing & Allied Health
Center for Biostatistics and Advanced Informatics
3901 Rainbow Blvd.
Kansas City, Ks 66160

Cynthia Teel, RN, PhD
Associate Professor, Assistant Dean, Academic Affairs
University of Kansas School of Nursing
3901 Rainbow Blvd.
Kansas City, Ks 66160

Lauren S. Aaronson, RN, PhD, FAAN
Professor
University of Kansas School of Nursing
3901 Rainbow Blvd.
Kansas City, Ks 66160

Corresponding author:
Byron Gajewski, PhD
Address for reprint requests:
3901 Rainbow Blvd.
MS 4043
Kansas City, Ks 66160
Phone: 913-588-1603
E-mail: bgajewski@kumc.edu

Acknowledgements: Appreciation is extended to the KU Nurses Alumni Association (University of Kansas) for its generosity in funding the original research project, which served as the observed data in this paper.

Keywords: Bayesian data analysis, statistical analysis, T'ai Chi

1 ABSTRACT

2 Background: Bayesian inference provides a formal framework for updating knowledge by
3 combining prior knowledge with current data. Over the past ten years, the Bayesian paradigm
4 has become a popular analytic tool in health research and this popularity is likely to increase in
5 the near future. While the nursing literature contains examples of Bayes' theorem applications to
6 clinical decision-making, it lacks an adequate introduction to Bayesian data analysis.

7 Approach: Bayesian data analysis is introduced through a fully Bayesian model for determining
8 the efficacy of T'ai Chi as an illustrative example. The mechanics of using Bayesian models to
9 combine prior knowledge, or data from previous studies, with observed data from one's current
10 study are discussed.

11 Results: The primary outcome in the illustrative example was physical function. Three prior
12 probability distributions (priors) were generated for physical function using data from a similar,
13 previous study found in the literature. Each prior was combined with the likelihood from
14 observed data in the current study to obtain a posterior probability distribution. In each case, the
15 posterior distribution showed that the probability that the control group is better than the T'ai Chi
16 treatment group was low.

17 Discussion: Bayesian analysis is a valid technique that allows the researcher to manage varying
18 amounts of data appropriately. As advancements in computer software continue, Bayesian
19 techniques will become more accessible. Researchers must educate themselves on applications
20 for Bayesian inference, as well as its methods and implications for future research.

Bayesian Data Analysis: Estimating the Efficacy of Tai Chi as a Case Study

1 Nurse investigators attempt to answer research questions first by exploring the state of
2 the science; then by conducting research and gathering data to address the question or
3 hypothesis; and finally by updating their opinion or understanding based on both previous
4 findings and current outcomes. However, while hypotheses may be derived from past research,
5 classical statistical analysis is limited because it uses data gathered from the current study only,
6 essentially ignoring statistical information and data from the literature. Bayesian inference
7 provides a formal framework for specifically including and combining information and outcomes
8 from the literature with data from one's current study in order to update knowledge about the
9 phenomenon under study (Gelman, Carlin, Stern, & Rubin, 2004).

11 Over the past ten years, the Bayesian paradigm has become a popular analytic tool in
12 health research (Ashby, 2006), and has been applied in many technological, scientific, policy,
13 and medical settings (e.g. Grunkemeier & Payne, 2002; Jonsson, Jonsson, Bois, & Marshall,
14 2007; Nixon, & Thompson, 2005; Troiani & Carlin, 2004, respectively). The popularity of
15 Bayesian data analysis will likely increase in the near future (Little, 2006). While nursing
16 literature contains examples of Bayes' theorem applications to clinical decision making (e.g.
17 Caelli, Downie, & Caelli, 2003; Lee, Abbott, & Johantgen, 2005; Rudy, Lucke, Whitman, &
18 Davidson, 2001), it lacks a satisfactory introduction to Bayesian data analysis. Using an
19 example from the first author's research (Carpenter, 2006), Bayesian data analysis is introduced
20 by applying a fully Bayesian model to evaluate the efficacy of a T'ai Chi exercise intervention.

Classical Versus Bayesian Statistics

22 Traditionally, nurse investigators employ classical statistical inference to interpret
23 quantitative results. Classical inference, which uses only observed data from one's current study,

1 does not allow for the direct involvement of prior knowledge. Conversely, while Bayesian
2 inference also is based on the observed data, it additionally provides a mechanism for formally
3 using prior information in the analysis. The Bayesian paradigm incorporates different points of
4 view, evaluating current data in the context of previous information, to achieve a new level of
5 understanding. Because the researcher makes decisions about what prior literature to include in
6 the analysis, some have criticized Bayesian techniques for introducing subjectivity into the
7 statistical analysis. However, classical methods also involve subjective choices in the form of
8 implicit researcher biases, decisions about sampling plans, and assumptions about randomness of
9 distributions (Bullard, 2001)—which also certainly influence outcomes. As compared to classical
10 methods, Bayesian inference acknowledges researcher subjectivity, making it explicit as a direct
11 means of dealing with uncertainty in scientific inference (Eddy, 2004).

12 The process of Bayesian data analysis involves summarizing the state of the knowledge
13 using a *prior* probability distribution and, after collecting new data, revising this knowledge with
14 a *posterior* probability distribution (Albert, 1997). Similar to classical statistical inference,
15 which uses theoretical sampling distributions, the observed data in a Bayesian analysis are
16 summarized with a *likelihood* function. Because researchers' interpretations of existing
17 knowledge may vary, one single prior distribution for any particular research question is viewed
18 as inadequate. Therefore, Bayesian methods use several prior distributions to determine effects
19 on the posterior distribution, thus allowing for alternative interpretations.

20 In the current example, three prior probability distributions, generated using data from a
21 previous T'ai Chi study (Song, Lee, Lam, & Bae, 2003), were identified to demonstrate and
22 compare the effects of these prior probability distributions on the posterior probability
23 distribution, given the observed data from the current study (Carpenter, 2006). Both the current

1 study (Carpenter, 2006) and the Song et al study used the same modified form of T'ai Chi (TC),
2 *T'ai Chi for Arthritis*. Using this Bayesian data analysis approach, a two-sample comparison
3 (Gönen, Johnson, Lu, & Westfall, 2005) between the T'ai Chi group and the control group was
4 conducted for the primary outcome variable (physical function).

5 The purpose of Carpenter's study (2006) was to evaluate the effectiveness of *TC for*
6 *Arthritis* in individuals with knee/hip OA, measured by the Western Ontario-McMaster
7 Universities OA 3.1 Index (WOMAC™ 3.1). The WOMAC™ 3.1 is a 24-item measure that
8 assesses three dimensions: self-reported physical function, pain, and stiffness in knee or hip OA
9 (5-point Likert-type response format). For the present example, we focus only on the physical
10 function outcome. Carpenter hypothesized that the TC group would demonstrate greater
11 improvements in physical function compared to the control group. Subjects included in the
12 analysis were 22 Caucasian adults (81% female; mean age = 81 years) with OA of the knee
13 and/or hip, residing in one of five Midwest retirement facilities. Review and approval by an
14 academic Health Center Human Subjects Committee was obtained prior to initiating the study.

15 The mean difference in physical function subscale scores (final minus baseline subscale
16 scores) was calculated for the TC group and for the control group. A negative mean difference
17 subscale score indicates improvement in physical function. Using classic data analysis, there
18 was a significant difference between the two groups, $t(20) = -2.31, p = 0.0158$ (one-tailed). The
19 TC group showed a mean decrease in physical function scores ($-5.00, SD 10.06, n = 12$),
20 suggesting improvement in physical function. In contrast, the control group showed a mean
21 increase in physical function scores ($4.64, SD 9.37, n = 10$), which suggests deterioration.

22 Traditionally, in classic data analysis the p -value says nothing about the probability of
23 either group performing better. The interpretation of the obtained $p = 0.0158$ is only the

1 probability of observing $t = -2.31$, or a more extreme value, given that the null hypothesis is true.
2 Based on this interpretation, one then decides whether or not to reject the null hypothesis. From
3 the Bayesian perspective, however, a probability interpretation of group performance is valid
4 (Gelman et al., 2004). Thus, a Bayesian approach may be more meaningful clinically as it
5 allows stating a probability for the comparison of physical function improvement between the
6 control group and the TC group, rather than only the traditional test against a null hypothesis.

7 Bayesian Data Analysis

8 To obtain the Bayesian interpretation of the effects of TC on physical function, a prior
9 probability distribution must first be determined. Representing background knowledge
10 (Goodman, 1999) the prior distribution is typically determined before the current data are
11 observed (Bullard, 2001). The effect that the prior distribution has on the posterior probability
12 distribution depends on its precision and on the precision of the current data. Alternative prior
13 distributions should be justified by published research from similar studies or by expert opinion
14 (e.g. O'Hagan et al., 2006). In the present example, a single study examining the effects of the
15 same modified TC (Song et al., 2003) as that used in the current study was available to determine
16 a prior probability distribution. In the Song et al. study, the TC group demonstrated greater
17 improvements in WOMAC™ 3.1 scores for self-reported physical function (mean difference
18 from baseline to end of study was -11.09 , SD 12.0, $n = 22$) compared to the control group (mean
19 difference of -1.33 , SD 10.06, $n = 21$). These data informed the prior probability distribution for
20 physical function (Figure 1).

21 Given possible debate about how to weight the Song et al. data, because of heterogeneity
22 across the two studies (Song et al. 2003; Carpenter, 2006), three priors were calculated to
23 demonstrate alternatives for weighting previous knowledge. In each of the three prior

1 distributions in Figure 1, Song et al. (2003) data were assigned different weights by using
2 differing amounts of prior information. Although it may seem counter-intuitive, the prior
3 distributions calculated are interpreted as the probability that the average improvement in
4 physical function for the control group was better than that for the TC group.

5 In the first *informative prior* (prior 1), the most weight was assigned to the data by using
6 means generated from the full Song et al. data set ($N=43$; $n=22$ in the TC group and $n=21$ in the
7 control group). Based on prior 1, the probability that the control group performed better than the
8 TC group on physical function is very low (probability= 0.0024). This probability is defined as
9 the area to the right of zero in Figure 1. In the second prior (or *semi-informative prior*: prior 2),
10 calculations from approximately one fourth of the Song, et al sample ($N=10.75$; $n=5.5$ in the TC
11 group and $n=5.25$ in the control group), was used in order to assign less weight to the prior study.
12 In this case, the probability that the control group performed better than the TC group on
13 physical function is increased (probability= 0.0792), but still low. Finally, a flat or *non-*
14 *informative prior* (prior 3), which assumes very little previous knowledge and is uniform across
15 all possible values of a parameter (Bullard, 2001), was generated. A non-informative prior
16 assigns very little weight to the previous study. In this case, the Song et al. (2003) data were
17 weighted much less than in the first two priors by using even less data ($N=2$; $n=1$ in the TC
18 group and $n=1$ in the control group). With this limited information, the probability that the
19 control group performed better than the TC group increased considerably (probability= 0.500). In
20 other words, the non-informative prior assigned a 50-50 probability that the control group
21 performed better than the TC group.

22 After identifying the three alternative priors, data from the current study (Carpenter,
23 2006), the observed data, were used to compute a single likelihood function (Figure 2). This

1 likelihood was calculated as a function of the mean difference between the TC and control group
2 in the Carpenter study. Next, by combining the prior distribution (from the Song et al. 2003
3 data) with the likelihood or observed data (from the Carpenter data), a posterior probability is
4 generated. The posterior probability distribution is defined as updated knowledge and it, in turn,
5 may serve as a prediction (or prior) for future studies.

6 In the present example, a posterior distribution was computed by combining each of the
7 three prior distributions with the likelihood function for the observed data (Figure 3). The
8 informative prior (prior 1) combined with the likelihood function resulted in a posterior or
9 updated probability of 0.0001 that the control group will perform better than the TC group. This
10 posterior probability distribution is lower than the original informative prior distribution
11 (0.0024), showing added information from the current study. That is, by including data from
12 both studies one obtains an even lower probability that the control group will perform better than
13 the TC group than one obtained with the first study alone (i.e., the informative prior). Likewise,
14 combining the semi-informative prior (prior 2) and the likelihood function from the current study
15 resulted in a posterior probability of 0.0034 that the control group will perform better than the
16 TC group--much lower than the original semi-informative prior (0.0792). Finally, the non-
17 informative prior (prior 3), when combined with the likelihood function, generated a posterior
18 probability of 0.0105, an even greater difference and a substantially lower probability than the
19 non-informative prior (0.500).

20 It is important to note that while the posterior probability distribution generated from the
21 non-informative prior appears to be similar to the traditional p -value obtained in the original
22 study (0.0158), its interpretation is quite different. As opposed to the interpretation of the p -
23 value as the probability of obtaining the found outcome (or one more extreme) if the true

1 outcome was zero, the interpretation of the posterior probability is the probability that the control
2 group will perform better than the TC group, which in this case is very small (0.0105), and thus
3 offers a more meaningful interpretation than that of a traditional p -value.

4 Discussion

5 Bayesian data analysis generated three prior probability distributions, which suggested
6 the control group was not as efficacious as the TC group at improving physical function. All
7 three posterior probability distributions, calculated by combining each prior from the Song et al.
8 data with the likelihood function from the current study data, indicated greater evidence that the
9 control group did not perform as well on physical function as the TC group than that obtained
10 from the original traditional analytic approach using a t -test. This is because all three posterior
11 distributions supported the conclusion that the control group did not perform as well as the TC
12 group. It is possible, however, that Bayesian analysis conducted on future similar studies with
13 larger sample sizes, using results from Song et al. (2003) and Carpenter (2006) as prior data,
14 might yield different posterior distributions and lead to different conclusions. In Bayesian terms,
15 results from each study investigating the same standardized TC for older adults with OA gives
16 subsequent researchers more information to consider for their prior distributions. This is the
17 particular strength of Bayesian analysis for building the evidence needed for translation of
18 research into practice.

19 There are, however, limitations to using Bayesian techniques in data analysis. As with
20 any statistical technique, models developed using Bayesian analysis are dependent on the quality
21 of the observed data. Unique to Bayesian analysis, the quality of prior knowledge, or data,
22 influences the model as well. Fortunately, sensitivity analysis across various priors helps to
23 minimize limitations. Although Bayesian analysis can be conducted with any sample size, the

1 more data collected in a study (i.e., a larger the N in the current study), the less influence the
2 prior information or data will have on the posterior distribution (Bullard, 2001). Also, because
3 goodness of fit is an important aspect of any parametric analysis, it is necessary to check the
4 assumption of normal distribution (Gelman, Meng, & Stern, 1996). For analytic purposes, it is
5 assumed that the prior distribution and the sampling distribution of the observed physical
6 function data are all in the family of normal distributions. When both have normal distributions,
7 the posterior distribution is also normal (Gelman et al., 2004). Finally, in the present example,
8 only a single outcome variable was used to introduce Bayesian methods. This may misrepresent
9 the complexity of the process as Bayesian analysis can be extremely complex. For examples of
10 more complex Bayesian data analyses involving nursing home research, see Gajewski et al.
11 (2006) or Gajewski, Thompson, Dunton, Becker, & Wrona (2006).

12 Conclusion

13
14 Bayesian analysis enables the researcher to consider each data set within the context of
15 previous knowledge, and to determine how the current data adds to or alters the researcher's
16 belief about the model (Goldstein, 2006). On the other hand, traditional statistical analysis relies
17 on theoretical sampling distributions of the data (Lee & Wagenmakers, 2005) and treats every
18 data set as if no context exists from which to consider the data (Goldstein). Bayesian analysis is a
19 valid technique that allows the researcher to manage varying amounts of data appropriately and
20 to formally incorporate previous findings in the analysis of data from one's current study—thus
21 strengthening conclusions drawn. As advancements in computer software continue, Bayesian
22 techniques will become more accessible. Researchers must educate themselves on the
23 applications for Bayesian inference, as well as its methods and implications for future research..

References

- Albert, J. (1997). Teaching Bayes' rule: A data-oriented approach. *The American Statistician*, 51(3), 247-253.
- Ashby, D. (2006). Bayesian statistics in medicine: A 25 year review. *Statistics in Medicine*, 25(21), 3589-3631.
- Bullard, F. (2001). *A brief introduction to Bayesian statistics*. Retrieved March 15, 2007 from North Carolina School of Science and Mathematics: <http://courses.ncssm.edu/math/TALKS/PDFS/BullardNCTM2001.pdf>
- Caelli, K., Downie, J., & Caelli, T. (2003). Towards a decision support system for health promotion in nursing. *Journal of Advanced Nursing*, 43(2), 170-180.
- Carpenter, J. (2006). *Effects of modified t'ai chi exercise on physical function, pain, and stiffness in older adults with lower extremity osteoarthritis* Unpublished Dissertation, University of Kansas Medical Center, Kansas City.
- Eddy, S. (2004). What is Bayesian statistics? *Nature Biotechnology*, 22(9), 1177-1178.
- Gajewski, B. J., Thompson, S., Dunton, N., Becker, A., & Wrona, M. (2006). Inter-rater reliability of nursing home surveys: A Bayesian latent class approach. *Statistics in Medicine*, 25(2), 325-344.
- Gajewski, B. J., Lee, R., Thompson, S., Dunton, N., Becker, A., & Wells, V. (2006). Non-normal path analysis in the presence of measurement error and missing data: A Bayesian analysis of nursing homes' structure and outcomes. *Statistics in Medicine*, 25(21), 3632-3647.
- Gelman, A., Meng, X.L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies, with discussion. *Statistica Sinica*, 6, 733-807.

- Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B. (2004). *Bayesian data analysis* (2nd ed.). Washington DC: Chapman and Hall/CRC.
- Goldstein, M. (2006). Subjective Bayesian analysis: Principles and practice. *Bayesian Analysis*, 1(3), 403-420.
- Gönen, M., Johnson, W.O., Lu, Y., & Westfall, P.H. (2005). The Bayesian two-sample t test. *The American Statistician*, 59(3), 252-257.
- Goodman, S. (1999). Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of Internal Medicine*, 130(12), 1005-1013.
- Grunkemeier, G., & Payne, N. (2002). Bayesian analysis - A new statistical paradigm for new technology. *The Annals of Thoracic Surgery*, 74(6), 1901-1980.
- Jonsson, F., Jonsson, E., Bois, F., & Marshall, S. (2007). The application of a Bayesian approach to the analysis of a complex, mechanistically based model. *Journal of Biopharmaceutical Statistics*, 17(1), 65-92.
- Lee, S. M., Abbott, P., & Johantgen, M. (2005). Logistic regression and Bayesian networks to study outcomes using large data sets. *Nursing Research*, 54(2), 133-138.
- Lee, M., & Wagenmakers, E.J. (2005). Bayesian statistical inference in psychology: Comment on Trafimow (2003). *Psychological Review*, 112(3), 662-668.
- Little, R. J. (2006). Calibrated Bayes: A Bayes/frequentist roadmap. *The American Statistician*, 60(3), 213-223.
- Nixon, R., & Thompson, S. (2005). Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. *Health Economics*, 14(12), 1217-1229.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., et al.

(2006). *Uncertain judgments: Eliciting experts' probabilities*. New York: Wiley.

Rudy, E., Lucke, J., Whitman, G., & Davidson, L. (2001). Benchmarking patient outcomes.

Journal of Nursing Scholarship, 33(2), 185-189.

Song, R., Lee, E., Lam, P., & Bae, S. (2003). Effects of tai chi exercise on pain, balance, muscle strength, and perceived difficulties in physical functioning in older women with

osteoarthritis: A randomized clinical trial. *Journal of Rheumatology*, 30, 2039-2044.

Troiani, J., & Carlin, B. (2004). Comparison of Bayesian, classical, and heuristic approaches in

identifying acute disease events in lung transplant recipients. *Statistics in Medicine*, 23, 803-824.

Figure Legends

Figure 1. Prior Probability Distributions

Figure 2. Likelihood Function

Figure 3. Posterior Probability Distributions

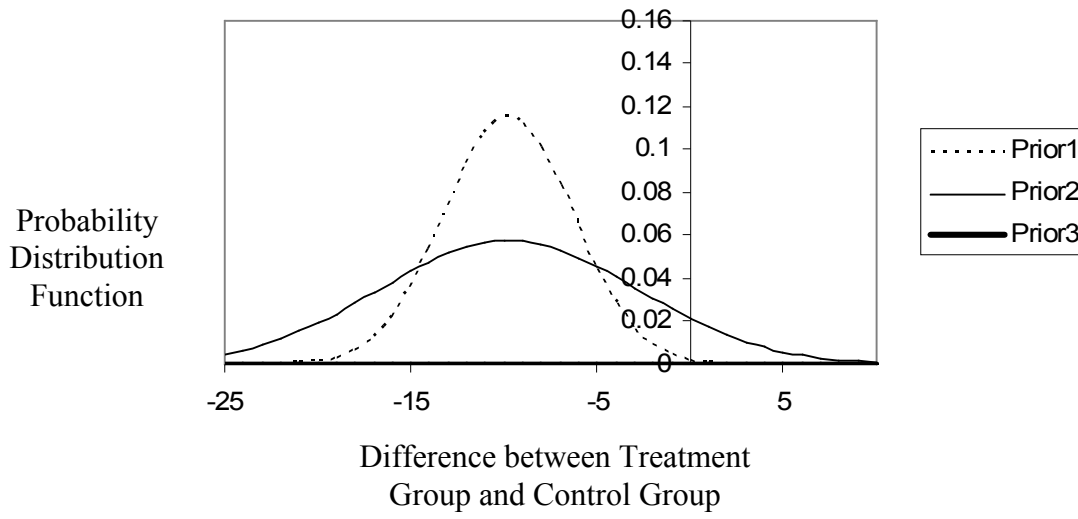


Figure 1. Prior Probability Distributions

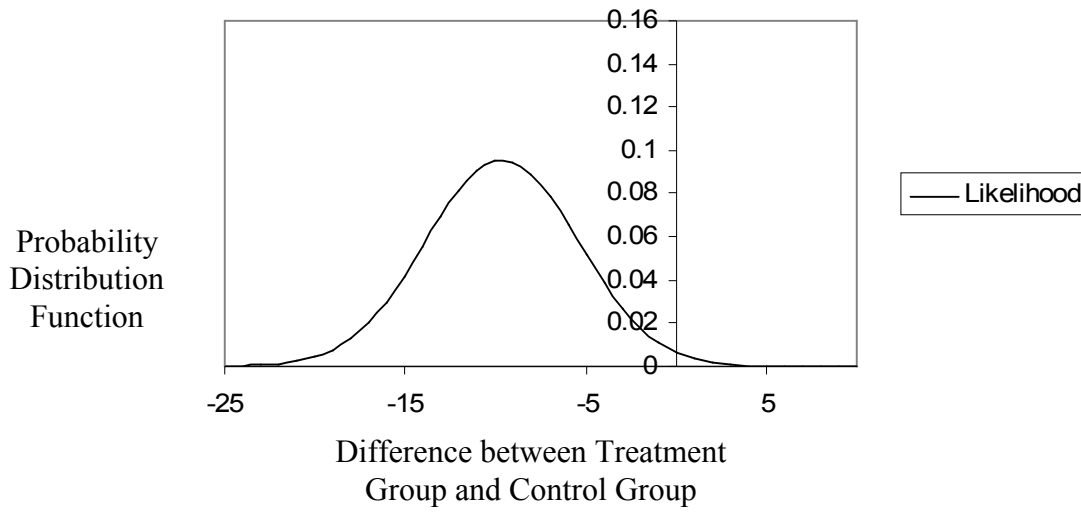


Figure 2. Likelihood Function

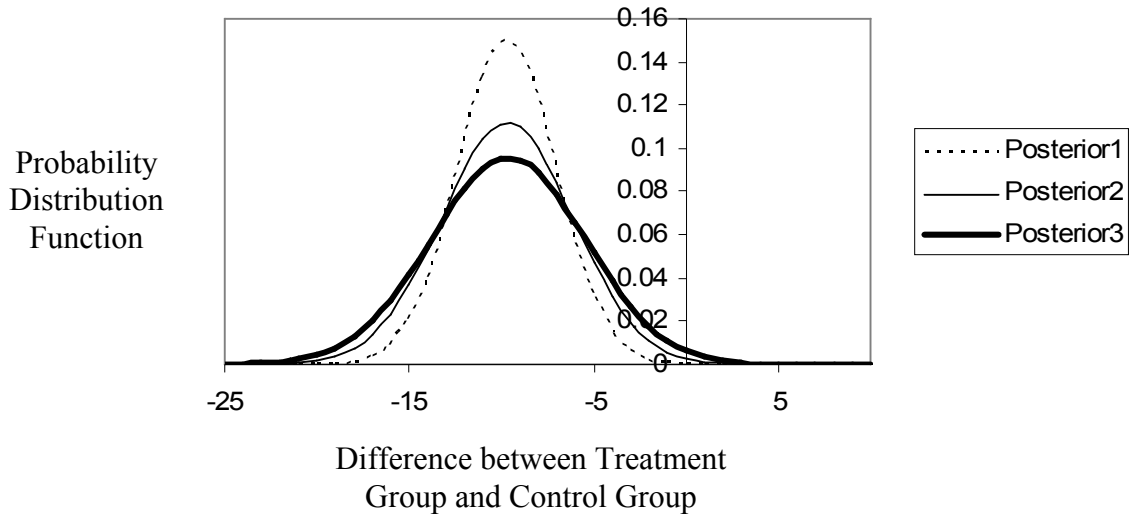


Figure 3. Posterior Probability Distributions