SPRINGER BRIEFS IN STATISTICS

Jeff Grover

Strategic Economic Decision-Making

Using Bayesian Belief Networks to Solve Complex Problems

Springer

SpringerBriefs in Statistics

For further volumes:
http://www.springer.com/series/8921

Jeff Grover

Strategic Economic Decision-Making

Using Bayesian Belief Networks to Solve Complex Problems

Springer

Jeff Grover
Maple Crest Way 512
Elizabethtown, Kentucky, USA

ISSN 2191-544X ISSN 2191-5458 (electronic)
ISBN 978-1-4614-6039-8 ISBN 978-1-4614-6040-4 (eBook)
DOI 10.1007/978-1-4614-6040-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012951657

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Sergeant First Class Charles V. Lang, IV
(U.S. Army, Retired)
"Friend"

Contents

Chapter 1
An Introduction to Bayes' Theorem and Bayesian Belief Networks (BBN)

1.1 Introduction to Bayes' Theorem and BBN

Determining future states of nature based on complex streams of asymmetric information is increasing comes at a premium cost for today's organizations across a global economy. Strategic leaders at all levels face uncanny events where information feeds at near real-time require decision-making based on the interactive effects of this information and across all spectrums of operations, to include the militaries, governments, corporations, and the scientific communities. The dominate information that has historically been absent here is subjective in nature and flows directly from the innate knowledge of leaders and subject matter experts (SME) of these organizations. With the use of inductive reasoning, we can integrate this truth and have a more plausible future expectation based on the decisions these leaders make today when we filter it through the lens of Bayes' theorem. This is done by formulating a hypothesis (a cause) of the proportional relationships one believes that exists and then filtering this knowledge through observable information (the effect(s)) to revise the initial beliefs.

There is a gradual acceptance by the scientific community of traditionalists (frequentists) for the Bayesian methodology. This is not through any new theoretical revelation, but through the sheer momentum of its current utility in scientific discovery. It possesses the uncanny ability to allow researchers to seamlessly transition from the traditional cause and effect to the effect and cause scenario using inductive logic or plausible reasoning.[1] This precipice is possible, in part, through the use of subjective (prior) beliefs where researchers obtain knowledge, either through historical information or subject matter expertise, when attempting

[1] E. T. Jaynes, in his book, "*Probability Theory: The Logic of Science*" (Jaynes 1995) suggests the concept of plausible reasoning is a limited form of deductive logic and "The theory of plausible reasoning" ... "is not a weakened form of logic; it is an extension of logic with new content not present at all in conventional deductive logic" (p. 8).

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_1,

to formulate the truth. This knowledge[2] can originate either from observed data or intuitive facts as seen through the lenses of these SME. The use of this "prior" knowledge, though, manifests the theoretical rub or clash between the traditionalists and Bayesians. The question remains, even in the midst of this struggle; does the Bayesian method have utility?

In responding to this question, I can offer an example from the art and science of diagnosing a disease. There is a great consensus that the science of diagnosing is unequivocally rigid and more exact. The art of diagnosing though is not an exact science. Let's consider a case and solve it using the basic concepts of Bayes' theorem. Suppose you are to undergo a medical test to rule out a horrific disease but test result is positive (Event A), which suggests that you have the disease (Event B). We are conditioning Event A on B and we are expressing the relationship, which says, the probability of the test being positive, P(A), given you have the disease P(B), or $P(A|B)$. What we are looking for the opposite—the probability that you have the disease given the test is positive, $P(B|A)$. If we let P(A) = 5.9 %,[3] P(B) = 1 %, and $P(A|B) = 95.0\,\%$, then we have enough information to answer $P(B|A)$. Using Set theory, we are interested in the sharing of these two random events going in both directions. First, we are interested in $P(A \cap B)$ and then $P(B \cap A)$. Knowing that $P(A \cap B = P(B \cap A)$ we can use the chain rule of probability to get our answer. Since $P(A \cap B) = P(A)\ P(B|A)$ and $P(B \cap A) = P(B)\ P(A|B)$, we have $P(A)\ P(B|A) = P(B)\ P(A|B)$. Rearranging we have for one path:

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

Now, we can simply solve this equation and obtain our answer.

$$P(B|A) = \frac{\mathrm{P}(1.0\,\%)\ \mathrm{P}(95.0\,\%)}{\mathrm{P}(5.9\,\%)} = \frac{1.0\,\%}{5.9\,\%} = 16.1\,\%.$$

Now, the probability of actually having the disease given you have a positive test result is downgraded from 95 % to 16.1 %, which is a non-significant percentage compared to just chance alone. This is the essence of Bayes' theorem—it has the ability to slice through observable information using prior beliefs to reweight the truth proportionally.

Now, consider the scenario where your physician makes a diagnosis based on the test results of 95.0 % and recommends surgery, a regiment of medicine, or even additional tests. If she or he is incorrect in their diagnosis, then the economic consequences at a minimum would include the psychological costs of mental,

---

[2] In the BBN literature, researchers refer to this knowledge as subjective or originating from a priori (prior) probabilities.

[3] I computed the (marginal) probability of Event B, P(A), as $P(B) \times P(A|B) + P(\tilde{B}) \times P(A|\tilde{B}) = 1.0\,\% \times 95.0\,\% + 99.0\,\% \times 5.0\,\% = 0.95\,\% + 4.95\,\% = 5.9\,\%$.

physical and emotional pain and suffering but also those costs associated with surgery and a regiment of medication, where often the cure is worse than the cause.

When your physician begins to add prior knowledge or initial beliefs to this case, the original diagnosis comes into question. Suppose that only 1.0 % of the population actually has this disease? Using inductive logic, your physician would begin to adjust her or his beliefs of the diagnosis downward. Again, what if your physician adds the fact that there is no family history? Then she or he would continue to adjust their beliefs, possibly to non-significant levels. Doing this, manifest the underlying principles of BBN; they learn from these partial truths or knowledge.[4] Deductively, you would be less confident in the initial diagnosis if your physician did not consider these initial facts when making their diagnosis.

1.2 The Identification of the Truth

The above discussion begs for a discussion on the definition of truth. This is a critical discussion in the study of Bayes' theorem because if we accept that prior knowledge has intrinsic value, then we are well on our way to the use of BBN. If we reject this form of knowledge, then by default we remain in the traditionalist camp; and if we accept this form of knowledge, then we enter the Bayesian camp.

The latter accepts this knowledge as truth based on initial assumptions of rational beliefs. The former initially rejects this knowledge as truth through falsification using the rigors of hypothesis testing. The strength in the latter is that in searching for the truth, they accept partial truths, which begins to illuminate it more clearly. The strength in the former is their belief in the rigor of science. The facts remain—both parties have a dim view of the truth, initially, and each attempts to discover it logically deductively and inductively, respectively. For if we knew the truth, then there would be no need to search

The common Bible verse references the truth: "You will know the truth, and the truth will set you free" (John 8:32, New International Version). We are on a quest to find the "Golden Grail" of truth. If we knew the truth, then there would be no need to search. Since, generally speaking, we do not know the truth; we search for it using available information or knowledge. This comes from rigid systematic research and discoveries and from historical or innate knowledge hidden in traditions and facts known only to SME. Inductive logic (Bayesian logic) is allowing the scientific community to overcome traditional constraints induced by using the deductive logic of falsification. We are now able to overcome this gap by accepting the truth using Bayes' theorem.

[4] I use information and knowledge interchangeably throughout the book.

1.3 The Motivation for This Book

Firstly, I am motivated to share the concept of Bayes' theorem due to its simplicity and utility in everyday life. It has an uncanny ability to separate truth from fiction and can truthify information very rigidly, logically, and quickly. As we experience everyday life events with assumed facts presented as truths, we can apply inductive logic and inverse probability to back out of this noisy information and point toward the truth. The global economy continually bombards today's scientific, political, religious, government, marketing, and business world with vast amounts of information from terabytes of data in the form of information. We are motivated in finding the truth contained in this information, and to inductive logic of Bayes' theorem will provide a filter to allow us to see it more clearly. Secondly, I am also motivated to share this concept with strategic decision-makers in that Bayes' theorem is very robust in its ability to absorb SME expertise, without the need to use real data, where these decision makers can make plausible assumptions from this hidden information. Also, as their respective governmental, economic, and academic environments bombard with these data, knowing the question to ask empowers them to require the truth from both structured and unstructured data sources. With this power, they now have the uncanny ability to make decisions using all possible sources of information with a refined agreement of the truth contained therein.

1.4 The Intent of This Book

I intend to present the elementary principles of Bayes' theorem using minimal statistical terminology and symbology to allow for non-statisticians and naïve learners to learn quickly and apply these benefits symmetrically and seamlessly when modeling a BBN, which is absent from the literature. This is the literature contribution of this book to the study of Bayes' theorem and BBN. The difficulty is learning this material is partly because of the literature, which is flooded with resources in learning BBN but requires an exponential learning curve to grasp due to its complex nature and scattered and haphazard statistical symbology. I also intend to bridge this gap by providing grounded constructs within BBN with multiple examples across areas of interest I previously suggested. Having a fundamental understanding of these constructs is essential in absorbing the concepts embedded in Bayes' theorem. In addition, I intend to provide the learner with the appropriate starter statistical concepts, terminology, and definitions and with a series of ten examples, ending with a two stage, 3-Node BBN to illustrate the concepts I put forth in this book.

1.5 The Utility of Bayes' Theorem

As I suggested above, the utility of Bayes' theorem reaches across all branches of science. While very simple in design, it requires sound inductive logic when applying independent and dependent causal relationships. Based on the results of past events, we are motivation to determine logical events that will affect a universe of partial truths to allow the illumination of the truth. What has great utility in the Bayesian universe is its utility to contain an infinite number of illuminating events that when invoked, it scales down the original set of events so that the resulting universe begins to learn the truth. Incredible, we begin with a subjective view of what we believe the truth of a universe holds and then by invoking multiple events, the truth begins to reveal itself. This is diametrically different from the philosophy of deductive logic of falsifying and never really accepting the truth.

1.6 Inductive Verses Deductive Logic

As we battle the forces of traditionalism and realism, the frequentists argue that deductive reasoning is the only way to the truth, and the Bayesian argues that the past reveals the truth, inductively. Of course, the former will immediately suggest that the latter is biasing their data selection process by reaching back to historical or past observable events to determining future states of nature, they remain in a theoretical rut by not illuminating the truth using subjective information. While the Bayes' are exponentially exploiting the universe of truth by doing this reach-back and suggesting reasonable future states of nature. Just ask Microsoft Corporation or Google in their use of Bayesian inference within their search engines, or ask the medical community when they correctly diagnose the existence or non-existence of cancer. Clearly, there is a utility in Bayes' theorem and the use of inductive logic.

In defining inductive and deductive logic, Bolstad (2007) suggests the former uses plausible reasoning to infer the truth contained within a statement to gauge the truth or falsehood of other statements that are consequences of the initial one. He also suggests that inductive logic goes from the specific to the general using statistical inferences of a parameter using observable data from a sample distribution. In addition, he suggests that deductive logic proceeds from the general to the specific to infer the truth of a statement from knowing the truth or falsehood from other statements that are consequences of the initial statement. Here, we make deductions from a population distribution rather than a parameter to determine the sampling distribution of a statistic. Furthermore, he suggests that when we have some event that has no deductive arguments available, we may use inductive reasoning to measure its plausibility by working from the particular to the general. He agrees with Richard Threlkeld Cox's (1946) sentiment that any set of plausibilities that satisfy these desired properties must operate according to the same rules of probability. Now, we can logically revise plausibilities by using the rules of probability, which allows the use of prior truths to project future states of nature.

1.7 Popper's Logic of Scientific Discovery

Karl Popper, the father of deductive scientific reasoning, basically rejects inductive reasoning. For example, he asserts that just because we always see white swans, that does not mean that there or non-white ones. He believes only truths can be falsified or rejected, i.e., the rejection of the null in classical statistical. Here is where the Bayesians continue to conflict with the current scientific status quo as put forth by his idea of falsification and rejection of inductive reasoning. He asserts that we cannot prove but only disprove or falsify the truth. Bayes' updating of prior probability through iteratively invoking partial truth is in direct opposition of his assertion. In current scientific hypothesis testing, we only reject or fail to reject the null-we never prove it as absolute truth. Here, we go away from the truth, whereas with Bayes', we go towards it.

1.8 Frequentist Verses Bayesian (Subjective) Views

Following a Google search for the terms frequentist and Bayesian (subjective), representing the two schools of statistical thought, I quickly noted the interest in the latter. This search produced 174,000 results for frequentist and 4,980,000 for Bayesian. The latter etymology begins with M. G. Kendall who first used the term to contrast with Bayesians, whom he called "non-frequentists" (Kendall 1949). Given the difference between these two schools of thought, I will provide some discussion that will differentiate between them and provide insights to the philosophy that substantiates them. These dominant thoughts caused the frequentist view to overshadow the Bayesian during the first half of the twentieth century. We see the word "Bayesian" appear in the 1950s and by the 1960s, it became the term preferred by people who sought to escape the limitations and inconsistencies of the frequentist approach to probability theory.

1.8.1 Frequentist to Subjectivist Philosophy

John Maynard Keynes(1921) provides a treatise on the role of the frequentist. His chapter VIII, "*The Frequency Theory of Probability*" provides 17 points of insight to the position of subjectivism, which follow:

- Point 1 suggests the difficulty in comparing degrees of probability of the frequentist and offers an alternative theory.
- Point 2 suggests a link to frequentist theory back to Aristotle who stated that: "the probability is that which for the most part happens" (p. 92). Keynes traces the frequentist back to Leslie Ellis who he suggested invented the concept that: "If the probability of a given event be correctly determined" … "the event will on a long run of trails tend to recur with frequency proportional to their probability" (p. 93). He also suggests that Venn, in his "*Logic of Chance*" was an early adopter.
- Point 3 suggests that Venn expresses an interest in probabilities through an empirically determined series of events and suggested that one may express probabilities based on experience.
- Point 4 suggests a divergence of probability from frequentist statistics, as initiated by Venn.
- Point 5 suggests that Venn's theory is narrowly limited in his exclusion of events that are not certain from the science of probability, which allows us to express statements of frequency. Venn also suggests that these probabilities can be derived either through inductive or deductive logic.
- Point 6 suggests two points where we have "induced Venn to regard judgments based on statistical frequency" (p. 97) into the frequentist camp are subjectivity and the inability for us to provide accurate measurements, Venn fails to discuss these in his theory. So, they are not ruled out, if you will, as being subjective in nature.
- Point 7 suggests then that Venn's theory is incomplete because he admits that in most cases we can arrive at statistical frequencies using induction.
- Point 8 suggests that Venn's belief was that we base probabilities on statistical frequencies alone, which are based on calculable chance. Most importantly, Keynes brings to the discussion the concept of inverse and *a posteriori* probabilities based on statistical grounds.
- Point 9 suggests that Karl Pearson agrees with Venn but only generally. He suggests a generalized frequency theory that does not regard probability to be identical with statistical frequency.
- Point 10 suggests the use of true proportions as a class of true frequencies as the measure of the probability of a proportion relative to a class, which is equal to the truth-frequency. Alternatively, that "the probability of a proportion always depends upon referring it to some class whose truth-frequency is known within wide or narrow limits" (p. 101). This gives rise to the idea of conditional probability, which gives probabilities of proportions that are relative to given data.
- Point 11 suggests criticism of frequency theory based on how one determines the class of reference, which we cannot define as "being the class of proportions of which everything is true is known to be true of the proportion whose probabilities we seek to determine" (p. 103).
- Point 12 suggests a modified view of frequency theory based on the above argument.

- Point 13 suggests the "Additional Theorem" (p. 105) which is based on how to derive true proportions that are "independent for knowledge" (p. 106) relative to the given data. This points us to a theorem of "inverse probability" and the use of a-priori knowledge.
- Point 14 suggests one can base his theory of inverse probability on inductive reasoning.
- Points 15–17 suggest additional arguments for his inverse theory of probability.

1.8.2 Bayesian Philosophy

While Bayes is the Father of Bayesian inference, we give credit to Pierre-Simon LaPlace for actually deriving the formula as we see it today. He transitioned probability science from the objective to the subjective school of thought. Here, the former purports that the statistical analysis depends only on the assumed model and the analyzed data, and that one did not require subjective decisions. Conversely, the subjectivist school did not require objective analysis for hypothesis determination. Fine (2004) reviews Joyce (2008) who suggests that Bayesian probability interprets the concept of probability as knowledge-base or inductive measure instead of the frequentist view of an event's probability as the limit of its relative frequency in a large number of trials. From the Bayesian view, the literature presents two views that interpret states of knowledge: the objectivist and the subjectivist school. The former is an extension of Aristotelian logic, and for the latter, the state of knowledge corresponds to a personal belief. The dominant feature of the Bayesian view is that one can assign a probability to a hypothesis, which one cannot do as a frequentist. The basis of Bayes' theorem in its simplest form is its ability to revise previous information when one invokes it to determine unique event revised probabilities. This statement requires a rigorous alternative approach to probability theory. Its essence is its ability to account for observed information when updating the unobservable. Understanding this concept is fundamental to learning Bayes' theorem. Chapter 2, "*A Literature Review of Bayes' Theorem and Bayesian Belief Networks (BBN)*," will discuss the evolution of Bayes' theorem and BBN in some detail.

Acknowledgments: I would like to thank the following researchers for their contributions to this chapter: (1) Anna T. Cianciolo, Ph.D., (2) Stefan Conrady, (3) Major Tonya R. Tatum, (4) Mrs. Denise Vaught, MBA, and (5) Jeffrey S. Grover Jr. Dr. Cianciolo is an Assistant Professor with the Department of Medical Education, Southern Illinois University School of Medicine, http://www.siumed.edu/; Stefan Conrady is the Managing Partner of Conrady Science, www.conradyscience.com; Major Tatum is an Operations Research/Systems Analyst with the Mission and Recruiter Requirements Division, Assistant Chief of Staff, G2, U.S. Army Recruiting Command; Denise Vaught is the President of Denise Vaught & Associates, PLLC, http://www.denisevaught.com/ and is a Registered Nurse, Certified Rehabilitation Registered Nurse a Certified Case Manager; and Mr. Jeffrey S. Grover Jr. is a Senior and Economics Major at the University of Kentucky, Lexington.

References

Bolstad, W. M. (2007). *Introduction to Bayesian statistics* (2nd ed.). Hoboken, NJ: John Wiley.

Cox, R. T. (1946). Probability, frequency, and reasonable expectation. *American Journal of Physics, 14*, 1–13.

Fine, T. L. (2004). The "only acceptable approach" to probabilistic reasoning. In E. T. Jaynes (Ed.), *A book review of: Probability theory: The logic of science*. Cambridge, UK: Cambridge University Press. 2003, 758. *SIAM News, 37*(2).

Jaynes, E. T. (1995). *Probability theory: The logic of science*. http://shawnslayton.com/open/Probability%2520book/book.pdf. Accessed 26 June 2012.

Joyce, J. (2008). "*Bayes' theorem*", *The Stanford encyclopedia of philosophy* (Fall 2008 edn), E. N. Zalta (Ed.). http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/. Accessed 1 July 2012.

Kendall, M. G. (1949). On the reconciliation of theories of probability. *Biometrika (Biometrika Trust), 36*(1/2), 101–116. doi:10.1093/biomet/36.1-2.101.

Keynes, J. M. (1921). *A treatise on probability*: Macmillan and Co. universal digital library. collection: Universallibrary. Retrieved from http://archive.org/details/treatiseonprobab007528mbp. Accessed 15 June 2012.

Chapter 2
A Literature Review of Bayes' Theorem and Bayesian Belief Networks (BBN)

2.1 Introduction to the Bayes' Theorem Evolution[1]

The concept of the theorem begins with a series of publications beginning with the "*Doctrine of Chances*" by Abraham de Moivre during the period of 1718–1756 (Schneider 2005). Historians have named the theorem after the Reverend Thomas Bayes[2] (1702–1761), who studied how to compute a distribution for the parameter of a binomial distribution. His friend, Richard Price (1763), edited and presented the work in 1763, after his death, as "*An Essay towards solving a Problem in the Doctrine of Chances*" (Bayes and Price 1763). Of particular importance is his Proposition 9. Of greater importance is Bayes' original idea of using a "Starting Guess" for a parameter of interest. This ignites the science of inverse probability and the beginning of a new school of probability thought. We see the different schools linked to philosophical approaches such as "Classical" statistics from R.A. Fisher's p-values and Aris Spanos Jerzy Neyman's deductive hypothesis tests, or the Popperian view of science that an hypothesis is made, and then it is tested and can only be rejected or falsified, but never accepted (Lehmann 1995). The Bayesian epistemology[3] runs contrary to these schools of thought.

In 1774, Pierre-Simon LaPlace publishes his first version of inverse probability following his study of Moivre's "*Doctrine of Chance,*" presumably the 1756 version. His final rule was in the form we still use today:

$$P(C|E) = \frac{P(E|C)P_{prior}(C)}{\sum P(E|C')P_{prior}(C')},$$

[1] I obtained most of the facts on the evolution of Byes' theorem included in Section 2.1 from McGrayne (2011).

[2] There is still debate on the true author of Bayes' theorem. Some give the honor to Pierre-Simon LaPlace following his 1774 publication of a similar theorem.

[3] See Joyce's comments on Bayesian epistemology for a complete discussion (Joyce 2008).

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_2,

where $P(C|E)$ is the probability of a hypothesis C(Cause) given data or information, which is equal to the probability of new information, $P(E|C)$ times the prior information divided by the sum of the probabilities of the data of all possible hypotheses. In the late 1870s early 1880s, Charles Sanders Peirce championed frequency-based probability, which launches this stream of empirical thought. In 1881, George Chrystal challenges Laplace's idea of the theorem[4] and declares that the laws of inverse probability are dead. Towards the end of the seventeenth Century, we begin to see some utility of the theorem when the French mathematician and physicist Henri Poincare' invokes the theorem during the military trial of the Dreyfus affair of 1899 to prove the falsity of this accusation that Alfred Dreyfus, a French army officer and Jew, was a German spy.

2.1.1 Early 1900s

In 1918, Edward C. Molina, a New York City engineer and self-taught mathematician, uses the theorem to evaluate the economic value of automating the Bell telephone system with call data to adopt a cost-effective strategy to deal with this uncertainty to prevent a looming bankruptcy. Albert Wurts Whitney, a Berkley insurance mathematics expert, uses the theorem to establish a form of social insurance with optimized premiums.

2.1.2 1920s–1930s

In 1926, Sir Harold Jeffreys, the Father of modern Bayesian statistics, uses Bayes' Rule to infer that the Earth's core is liquid and Frank P. Ramsey, English mathematician and philosopher, suggests making decisions under uncertainly using personal beliefs and quantified through making a wager. In 1933, Andrey Kolmogorov, a Soviet mathematician, suggests the use of the theorem as a method of firing back at a German artillery bombardment of Moscow using Bertrand's Bayesian firing system. In 1936, Lowell J. Reed, a medical researcher at Johns Hopkins University, uses the theorem to determine the minimum amount of radiation required to cure cancer patients while causing the least amount of damage. In 1938, Erik Essen-Möller, Swedish professor of genetics and psychiatry, develops an index of probability for paternity testing that was mathematically equivalent to the theorem that was in use for 50 years until the advent of DNA testing. Finally, in 1939, Harold Jeffreys, a geologist, publishes his theory of probability that uses the theorem as the only method to conduct scientific experiments with subjective probabilities.

---

[4] Here after I refer to Bayes' theorem as "the theorem."

2.1.3 1940s–1950s

In 1941, Alan Mathison Turning, the father of the modern computer, invents a Bayesian system of Bankurismus using banded Banburg strips looking for "fits" using sequential analysis to break the German secret Enigma coding machine. In 1942, Kolmogorov introduces firing dispersion theory, which is a Bayesian scoring system using a 50-50 guess for aiming artillery, and Alan Turing invents Bayesian Turingismus to deduce the patterns of cams surrounding the Tunny-Lorenz machine by using "gut feels" as prior probabilities. From 1943 to 1944, Max Newman, a British mathematician and code-breaker, invents the Colossus I and II machines and intercepts a message that Hitler gave Rommel ordering a delay of his attack in Normandy. In 1945, Ian Cassels, Jimmy Whitworth, and Edward Simpson, cryptanalysts use the theorem to Break Japanese code during WWII and John Gleason, a cryptanalyst, uses the theorem to break Russian code during the Cold War era. In 1947, Arthur L. Baily, an insurance actuary, resurrects Bayes' theory and demands the legitimate use of prior probabilities making justification from the Bible referencing one's personal belief could make all things possible. In 1950, he reads his work on Credibility procedures during an actuarial society banquet, citing LaPlace's form of the theorem and the combination of prior knowledge with observed data. In 1951, Jerome Cornfield, a history major working at the National Institute of Health, uses the theorem to provide a solid theoretical link that smoking does cause cancer; allowing epidemiologists to link this disease with causes. In 1954, Jimmie Savage, a University of Chicago statistician, publishes his revolutionary book, the "Foundations of Statistics," which extends Frank Ramsey's attempt to use the theorem for making inferences and decision-making. In 1955, L.H. Longly-Cook, a chief actuary, predicts the first U.S. catastrophic aviation disaster of two planes colliding in mid-air, which allows insurance companies to raise rates prior to this event and Hans Bühlmann, a mathematics professor, extends Baily's Bayes' philosophy and publishes a general Bayesian theory of credibility. In 1958, Albert Madansky, a statistician, writes a summary to the RANDS Corps final report, "*On the Risk of an Accidental or Unauthorized Nuclear Detonation,*" suggesting a probability greater than zero that this event could occur. Finally, in 1959, Robert Osher Schlaifer, a Harvard University's statistician, publishes "*Probability and Statistics for Business Decisions, An Introduction to Managerial Economics under Uncertainty*," which was a first reference to endorse the theorem.

2.1.4 1960s–Mid 1980s

In 1960, Morris H. DeGroot, a practitioner, publishes the first international text on Bayesian decision theory, Frederick Mosteller, Harvard University professor, and David L. Wallace, University of Chicago statistician, evaluate the 12 unknown authors of the Federalist papers using the theorem and identify Madison as the correct author. John W. Tukey, a Princeton statistic's professor, predicts Nixon as the winner of the Nixon-Kennedy presidential elections for NBC using their mainframe computers and Bayesian like code. In 1961, Homer Warner, a pediatric heart surgeon, develops the first computerized program for diagnosis of diseases using the theorem and Robert Osher Schlaiter and Howard Raiffa, two Harvard University business professors, publish "*Applied Statistical Decision Theory,*" a classical work using the theorem that charters the future direction for Bayesian theory. In 1968, John Piña Craven, civilian chief scientist, and Frank A. Andrews, Navy Captain, (retired), use Bayesian search techniques to locate the sunken submarine, the *U.S.S. Scorpion*. In 1974, Norman Carl Rasmussen, a physicist and engineer, uses Raiffas' decision trees (Raiffa 2012) to weigh the risks of meltdowns in the nuclear-power industry for the U.S. Nuclear Regulatory Commission (NRC) (Fienberg 2008). The NRC halts his study due to his inclusion of the theorem but following the 1979 Three Mile Island incident; they resurrected it. In 1975, Lawrence D. Stone, a Daniel H. Wagner Associates employee, publishes "Theory of Optimal Search" using Bayesian techniques following his participation in locating the *U.S.S. Scorpion* and the NRC gives him an invitation to publish his findings. In 1976, Harry C. Andrews, a digital image processor, publishes his "*Digital Image Restoration*." This uses Bayes' inference to restore nuclear weapons testing images from activity at Los Alamos National Laboratories. Finally, in 1983, Teledyne Energy Systems uses hierarchical methods to estimate shuttle failure at 35:1 when NASA estimated it as 100,000:1; in 1986, the Challenger explodes.

2.2 BBN Evolution

In 1985, Judea Pearl, computer scientist, publishes the seminal work on BBN, "*Bayesian Networks: A Model of Self Activated Memory for Evidential Reasoning*" (Pearl 1985) to guide the direction of BBN using discrete random variables and distributions. The following empirical studies are representative of peer review extensions to his work from 2005 to the present as queried through the Social Science Citation Index Web of Science® (Reuters 2012).[5]

2.2.1 Financial Economics, Accounting, and Operational Risks

BBN studies in these areas include: gathering information in organizations (Calvo-Armengol and Beltran 2009); conducting Bayesian learning in social networks (Acemoglu et al. 2011); processing information (Zellner 2002); evaluating games

[5] Certain data included herein are derived from the Web of Science ® prepared by THOMSON REUTERS ®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS ® 2012. All rights reserved.

and economic behavior (Mannor and Shinikin 2008), and economic theory and market collapse (Gunay 2008); determining accounting errors (Christensen 2010); evaluating operational risk in financial institutions (Neil et al. 2009); and determining the valuation of contingent claims with mortality and interest rate risks using mathematics and computer modeling techniques (Jalen and Mamon 2009).

2.2.2 Safety, Accident Analysis, and Prevention

BBN studies in these areas include: studying epidemiology; environmental, human safety, injury, and in accidents, road design, and urban settings (DiMaggio and Li 2012), evaluating infant mortality, deprivation, and proximity to polluting industrial facilities (Padilla et al. 2011) and human-centered safety analysis of prospective road designs (Gregoriades et al. 2010); predicting real-time crashes on the basic freeway segments of urban expressways (Hossain and Muromachi 2012) and crash counts by severity (Ma et al. 2008); evaluating the effects of osteoporosis on injury risk in motor-vehicle crashes (Rupp et al. 2010); and workplace accidents caused by falls from a height (Martin et al. 2009).

2.2.3 Engineering and Safety

BBN studies in these areas include: incorporating organizational factors into probabilistic risk assessment of complex socio-technical systems (Mohaghegh et al. 2009); predicting workloads for improved design and reliability of complex systems (Gregoriades and Sutcliffe 2008b); evaluating a methodology for assessing transportation network terrorism risk with attacker and defender interactions (Murray-Tuite and Fei 2010); evaluating individual safety and health outcomes in the construction industry (McCabe et al. 2008); evaluating risk and assessment methodologies at the work sites (Marhavilas et al. 2011); quantifying schedule risk in construction projects (Luu et al. 2009); and studying emerging technologies that evaluated railroad transportation of dangerous goods (Verma 2011).

2.2.4 Risk Analysis

BBN studies in this area include: developing a practical framework for the construction of a biotracing model as it applied to salmonella in the pork slaughterchain (Smid et al. 2011); assessing and managing risks posed by emerging diseases (Walshe and Burgman 2010); identifying alternative methods for computing the sensitivity of complex surveillance systems (Hood et al. 2009); assessing uncertainty in fundamental assumptions and associated models for cancer risk assessment (Small 2008); modeling uncertainty using model performance data (Droguett and Mosleh 2008); using Bayesian temporal source attribution to evaluate foodborne zoonoses (Ranta et al. 2011); and developing of posterior probability models in risk-based integrity modeling (Thodi et al. 2010).

2.2.5 Ecology

BBN studies in this area include: studying marine ecology to evaluate integrated modeling tools to support risk-based decision-making in marine spatial management (Stelzenmuller et al. 2011); integrating fuzzy cognitive mapping in a livelihood vulnerability analysis (Murungweni et al. 2011); optimizing participatory water resources management in Spain (Zorrilla et al. 2010); negotiating participatory irrigation management in the Indian Himalayas (Saravanan 2010); evaluating feral cat management options (Loyd and DeVore 2010); conducting an integrated analysis of human impact on forest biodiversity in Latin America (Newton et al. 2009); and integrating biological, economic, and sociological knowledge to evaluate management plans for Baltic salmon (Levontin et al. 2011).

2.2.6 Human Behavior

BBN studies in this area include: evaluating psychological and psychiatric factors in decision-making on ambiguous stimuli such as prosody by subjects suffering from paranoid schizophrenia, alcohol dependence, and without psychiatric diagnosis (Fabianczyk 2011); studying substance use and misuse and addiction to estimate population prevalence from the Alcohol Use Disorders Identification Test scores (Foxcroft et al. 2009); evaluating the role of time and place in the modeling of substance abuse patterns following a mass trauma (Dimaggio et al. 2009); and affective disorders on applied non-adult dental age assessment methods in identifying skeletal remains (Heuze and Braga 2008).

2.2.7 Behavioral Sciences and Marketing

BBN studies in these areas include: (1) Behavioral Sciences: analyzing adaptive management and participatory systems (Smith et al. 2007); evaluating human behavior in the development of an interactive computer-based interface to support the discovery of individuals' mental representations and preferences in decisions problems as they relate to traveling behavior (Kusumastuti et al. 2011); determining semantic coherence (Fisher and Wolfe 2011); conducting a behavioral and brain science study to evaluate base rates in ordinary people (Laming 2007); evaluating the implications of natural sampling in base-rate tasks (Kleiter 2007) and evaluating a probabilistic approach to human reasoning as a précis of Bayesian rationality (Oaksford and Chater 2009); and conducting an environmental and behavioral study to model and measure individuals' mental representations of complex spatio-temporal decision problems (Arentze et al. 2008). (2) Marketing: evaluating marketplace behavior (Allenby 2012); modeling a decision-making aid for competitive intelligence and marketing analysts (Michaeli and Simon 2008); and investigating endogeneity bias in marketing (Liu et al. 2007).

2.2.8 Decision Support Systems (DSS) with Expert Systems (ES) and Applications, Information Sciences, Intelligent Data Analysis, Neuroimaging, Environmental Modeling and Software, and Industrial Ergonomics

BBN studies in these areas include: (1) DDS with ES and Applications: aiding the diagnosis of dementia (Mazzocco and Hussain 2012); determining customer churn analysis in the telecom industry of Turkey (Kisioglu and Topcu 2011); conducting a customer's perception risk analysis in new-product development (Tang et al. 2011); assessing critical success factors for military decision support (Louvieris et al. 2010); predicting tourism loyalty (Hsu et al. 2009); Korean box-office performance (Lee and Chang 2009); and using data mining techniques to detect fraudulent financial statements (Kirkos et al. 2007) and (Ngai et al. 2011). (2) Information Sciences: evaluating affectively intelligent and adaptive car interfaces work (Nasoz et al. 2010). (3) Intelligent Data Analysis: evaluating automatic term recognition (Wong et al. 2009) and a socio-technical approach to business process simulation (Gregoriades and Sutcliffe 2008a). (4) Neuroimaging: conducting multi-subject analyses with dynamic causal modeling (Kasess et al. 2010); (5) Environmental Modeling and Software: evaluating perceived effectiveness of environmental DDS in participatory planning using small groups of end-users (Inman et al. 2011) and modeling linked economic valuation and catchment (Kragt et al. 2011); and (6) Industrial Ergonomics: exploring diagnostic medicine using DDS (Lindgaard et al. 2009).

2.2.9 Cognitive Science

BBN studies in this area include: evaluating the role of coherence in multiple testimonies (Harris and Hahn 2009) and a learning diphone-based segmentation (Daland and Pierrehumbert 2011); evaluating the efficiency in learning and problem solving (Hoffman and Schraw 2010); evaluating the base rate scores of the Millon Clinical Multiaxial Inventory-III (Grove and Vrieze 2009); evaluating spatial proximity and the risk of psychopathology after a terrorist attack (DiMaggio et al. 2010); evaluating actuarial estimates of sexual recidivism risk (Donaldson and Wollert 2008); and evaluating poor diagnostic reliability with sexually violent predator evaluations (Wollert 2007).

2.2.10 Medical, Health, Dental, and Nursing

BBN studies in these areas include: (1) Medical: evaluating the risk of tuberculosis infection for individuals lost to follow-up (Martinez et al. 2008) and assessing differences between physicians' realized and anticipated gains from electronic health record adoption (Peterson et al. 2011); (2) Health: evaluating socioeconomic inequalities in mortality in Barcelona (Cano-Serral et al. 2009); estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity (Elliott et al. 2008); and facilitating uncertainty in economic evaluations of patient level data (McCarron et al. 2009); (3) Dental: combining surveillance and expert evidence of viral hemorrhagic septicemia freedom (Gustafson et al. 2010) and investigating dentists' and dental students' estimates of diagnostic probabilities (Chambers et al. 2010); (4) Nursing: evaluating affective disorders in postnatal depression screening (Milgrom et al. 2011); estimating coronary heart disease risk in asymptomatic adults (Boo et al. 2012); determining the efficacy of T'ai Chi (Carpenter et al. 2008); and evaluating diagnostic test efficacy (Replogle et al. 2009).

2.2.11 Environmental Studies

BBN studies in this area include: identifying potential compatibilities and conflicts between development and landscape conservation (McCloskey et al. 2011); evaluating longer-term mobility decisions (Oakil et al. 2011); assessing uncertainty in urban simulations (Sevcikova et al. 2007); modeling land-use decisions under conditions of uncertainty (Ma et al. 2007); determining the impact of demographic trends on future development patterns and the loss of open space in the California Mojave Desert (Gomben et al. 2012); determining a methodology to facilitate compliance with water quality regulations (Joseph et al. 2010); and using participatory object-oriented Bayesian networks and agro-economic models for groundwater management in Spain (Carmona et al. 2011).

2.2.12 Miscellaneous: Politics, Geriatrics, Space Policy, and Language and Speech

BBN studies in these areas include: (1) Politics: a study evaluated partisan bias and the Bayesian ideal in the study of public opinion (Bullock 2009); (2) Geriatrics: an evaluation of the accuracy of spirometry in diagnosing pulmonary restriction in elderly people,(Scarlata et al. 2009); (3) Space Policy: the value of information in methodological frontiers and new applications for realizing asocial benefit (Macauley and Laxminarayan 2010); and (4) Language and Speech: include quantified evidence in forensic authorship analysis (Grant 2007) and causal explanation and fact mutability in counterfactual reasoning (Dehghani et al. 2012).

2.3 Current Government and Commercial Users of BBN

The following list represents the utility in current business, government, and commercial users of BBN:

- Analyzing information system network risk (Staker 1999).
- Analyzing roadway safety measures (Schultz et al. 2011).
- Applications in land operations (Starr and Shi 2004).
- Building process improvement business cases (Linders 2009).
- Comparing public housing and housing voucher tenants (The U.S. Department of Housing and Urban Development) (Mast 2012).
- Conducting social network analysis (Koelle et al. n.d.).
- Conducting unified, flexible and adaptable analysis of misuses and anomalies in network intrusion detection and prevention systems (Bringas 2007).
- Designing food (Corney 2000).
- Evaluating the risk of erosion in peat soils (Aalders et al. 2011).
- Evaluating U.S. county poverty rates (The U.S. Census Bureau) (Asher and Fisher 2000).
- Executing cognitive social simulation from a document corpus (The Modeling, Virtual Environments, and Simulation Institute) (McKaughan et al. 2011).
- Identifying military clustering problem sets (BAE Systems) (Sebastiani et al. 1999).
- Identifying potential compatibilities and conflicts between development and landscape conservation (McCloskey et al. 2011).
- Improving Attrition Rates in the M1A1/M1A2 Master Gunner Course (U.S. Army) (Zimmerman et al. 2010).
- Investigating engineering design problems (The U.S. Department of Energy) (Swiler 2006).
- Investigating the relationships between environmental stressors and stream condition (Allan et al. 2012).
- Measuring the internal dosimetry of uranium isotopes (The Los Alamos National Laboratory) (Little et al. 2003).
- Measuring neighborhood quality with survey data (The U.S. Department of Housing and Urban Development) (Mast 2010).
- Modeling the reliability of search and rescue operations within United Kingdom Coastguard (maritime rescue) coordination centres (Norrington et al. 2008).

- Optimizing and parameter estimation in environmental management (Vans 1998).
- Predicting long-term shoreline change due to sea-level rise (The U.S. Geological Survey Data Series) (Gutierrez et al. 2011).
- Predicting the impacts of commercializing non-timber forest products on livelihoods ecology and society (Newton et al. 2006).
- Predicting the reliability of military vehicles (Neil et al. 2001).
- Ranking of datasets (U.S. Government) (Data.gov).
- Use in U.S. Government public policy and government settings including: city growth in the areas of census-taking and small area estimation, U.S. election night forecasting, U.S. Food and Drug Administration studies, assessing global climate change, and measuring potential declines in disability among the elderly (Fienberg 2011).

References

Aalders, I., Hough, R. L., & Towers, W. (2011). Risk of erosion in peat soils – an investigation using Bayesian belief networks. *Soil Use and Management, 27*(4), 538–549. doi:10.1111/j.1475-2743.2011.00359.x.

Acemoglu, D., Dahleh, M. A., Lobel, I., & Ozdaglar, A. (2011). Bayesian learning in social networks. *Review of Economic Studies, 78*(4), 1201–1236. doi:10.1093/restud/rdr004.

Allan, J. D., Yuan, L. L., Black, P., Stockton, T., Davies, P. E., Magierowski, R. H., et al. (2012). Investigating the relationships between environmental stressors and stream condition using Bayesian belief networks. *Freshwater Biology, 57*, 58–73. doi:10.1111/j.1365-2427.2011.02683.x.

Allenby, G. M. (2012). Modeling marketplace behavior. *Journal of the Academy of Marketing Science, 40*(1), 155–166. doi:10.1007/s11747-011-0280-3.

Arentze, T. A., Dellaert, B. G. C., & Timmermans, H. J. P. (2008). Modeling and measuring individuals' mental representations of complex spatio-temporal decision problems. *Environment and Behavior, 40*(6), 843–869. doi:10.1177/0013916507309994.

Asher, J., & Fisher, R. (2000). Alternative scaling parameter functions in a hierarchical Bayes model of U.S. county poverty rates. *Proceedings of the Survey Research Methods Section, ASA*.

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S. *Philosophical Transactions of the Royal Society of London, 53*, 370–418. doi:10.1098/rstl.1763.0053.

Boo, S., Waters, C. M., & Froelicher, E. S. (2012). Coronary heart disease risk estimation in asymptomatic adults. *Nursing Research, 61*(1), 66–69. doi:10.1097/NNR.0b013e31823b1429.

Bringas, P. G. (2007). Intensive use of Bayesian belief networks for the unified, flexible and adaptable analysis of misuses and anomalies in network intrusion detection and prevention systems. *Database and Expert Systems Applications, 2007. DEXA '07. 18th International Workshop on, 3–7 September 2007* (pp. 365–371). doi:10.1109/dexa.2007.38

Bullock, J. G. (2009). Partisan bias and the Bayesian ideal in the study of public opinion. *Journal of Politics, 71*(3), 1109–1124. doi:10.1017/s0022381609090914.

Calvo-Armengol, A., & Beltran, J. D. (2009). Information gathering in organizations: Equilibrium, welfare, and optimal network structure. *Journal of the European Economic Association, 7*(1), 116–161.

Cano-Serral, G., Azlor, E., Rodriguez-Sanz, M., Pasarin, M. I., Martinez, J. M., Puigpinos, R., et al. (2009). Socioeconomic inequalities in mortality in Barcelona: A study based on census tracts (MEDEA project). *Health & Place, 15*(1), 186–192. doi:10.1016/j.healthplace.2008.04.004.

Carmona, G., Varela-Ortega, C., & Bromley, J. (2011). The use of participatory object-oriented Bayesian networks and agro-economic models for groundwater management in Spain. *Water Resources Management, 25*(5), 1509–1524. doi:10.1007/s11269-010-9757-y.

Carpenter, J., Gajewski, B., Teel, C., & Aaronson, L. S. (2008). Bayesian data analysis: Estimating the efficacy of t'ai chi as a case study. *Nursing Research, 57*(3), 214–219. doi:10.1097/01.NNR.0000319495.59746.b8.

Chambers, D. W., Mirchel, R., & Lundergan, W. (2010). An investigation of dentists' and dental students' estimates of diagnostic probabilities. *The Journal of the American Dental Association, 141*(6), 656–666.

Christensen, J. (2010). Accounting errors and errors of accounting. *The Accounting Review, 85*(6), 1827–1838. doi:10.2308/accr.2010.85.6.1827.

Corney, D. (2000). Designing food with Bayesian belief networks. Springer, pp. 83–94

Daland, R., & Pierrehumbert, J. B. (2011). Learning diphone-based segmentation. *Cognitive Science, 35*(1), 119–155. doi:10.1111/j.1551-6709.2010.01160.x.

Data.gov FAQ. http://www.data.gov/faq/. Accessed 16 June 2012.

Dehghani, M., Iliev, R., & Kaufmann, S. (2012). Causal explanation and fact mutability in counterfactual reasoning. *Mind & Language, 27*(1), 55–85. doi:10.1111/j.1468-0017.2011.01435.x.

DiMaggio, C., & Li, G. H. (2012). Roadway characteristics and pediatric pedestrian injury. *Epidemiologic Reviews, 34*(1), 46–56. doi:10.1093/epirev/mxr021.

Dimaggio, C., Galea, S., & Vlahov, D. (2009). Bayesian hierarchical spatial modeling of substance abuse patterns following a mass Trauma: The role of time and place. *Substance Use & Misuse, 44*(12), 1725–1743. doi:10.3109/10826080902963399.

DiMaggio, C., Galea, S., & Emch, M. (2010). Spatial proximity and the risk of psychopathology after a terrorist attack. *Psychiatry Research, 176*(1), 55–61. doi:10.1016/j.psychres.2008.10.035.

Donaldson, T., & Wollert, R. (2008). A mathematical proof and example that Bayes's theorem is fundamental to actuarial estimates of sexual recidivism risk. *Sexual Abuse-a Journal of Research and Treatment, 20*(2), 206–217. doi:10.1177/1079063208317734.

Droguett, E. L., & Mosleh, A. (2008). Bayesian methodology for model uncertainty using model performance data. *Risk Analysis, 28*(5), 1457–1476. doi:10.1111/j.1539-6924.2008.01117.x.

Elliott, M. N., Fremont, A., Morrison, P. A., Pantoja, P., & Lurie, N. (2008). A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity. *Health Services Research, 43*(5), 1722–1736. doi:10.1111/j.1475-6773.2008.00854.x.

Fabianczyk, K. (2011). Decision making on ambiguous stimuli such as prosody by subjects suffering from paranoid schizophrenia, alcohol dependence, and without psychiatric diagnosis. *British Journal of Mathematical and Statistical Psychology, 64*(1), 53–68. doi:10.1348/000711010x492366.

Fienberg, S. E. (2008). The early statistical years. 1947–1967: A conversation with Howard Raiffa. *Statistical Science, 23*(1), 136–149. doi:10.1214/088342307000000104 [Institute of Mathematical Statistics].

Fienberg, S. E. (2011). *Bayesian models and methods in public policy and government settings.* http://arxiv.org/pdf/1108.2177.pdf. Accessed 26 June 2012.

Fisher, C. R., & Wolfe, C. R. (2011). Assessing semantic coherence in conditional probability estimates. *Behavior Research Methods, 43*(4), 999–1002. doi:10.3758/s13428-011-0099-3.

Foxcroft, D. R., Kypri, K., & Simonite, V. (2009). Bayes' theorem to estimate population prevalence from alcohol use disorders identification test (AUDIT) scores. *Addiction, 104*(7), 1132–1137. doi:10.1111/j.1360-0443.2009.02574.x.

Gomben, P., Lilieholm, R., & Gonzalez-Guillen, M. (2012). Impact of demographic trends on future development patterns and the loss of open space in the California Mojave desert. *Environmental Management, 49*(2), 305–324. doi:10.1007/s00267-011-9749-6.

Grant, T. (2007). Quantifying evidence in forensic authorship analysis. *International Journal of Speech Language and the Law, 14*(1), 1–25. doi:10.1558/ijsll.v14i1.1.

Gregoriades, A., & Sutcliffe, A. (2008a). A socio-technical approach to business process simulation. *Decision Support Systems, 45*(4), 1017–1030. doi:10.1016/j.dss.2008.04.003.

Gregoriades, A., & Sutcliffe, A. (2008b). Workload prediction for improved design and reliability of complex systems. *Reliability Engineering and System Safety, 93*(4), 530–549. doi:10.1016/j.ress.2007.02.001.

Gregoriades, A., Sutcliffe, A., Papageorgiou, G., & Louvieris, P. (2010). Human-centered safety analysis of prospective road designs. *IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans, 40*(2), 236–250. doi:10.1109/tsmca.2009.2037011.

Grove, W. M., & Vrieze, S. I. (2009). An exploration of the base rate scores of the Millon clinical multiaxial inventory-III. *Psychological Assessment, 21*(1), 57–67. doi:10.1037/a0014471.

Gunay, H. (2008). The role of externalities and information aggregation in market collapse. *Economic Theory, 35*(2), 367–379. doi:10.1007/s00199-006-0158-7.

Gustafson, L., Klotins, K., Tomlinson, S., Karreman, G., Cameron, A., Wagner, B., et al. (2010). Combining surveillance and expert evidence of viral hemorrhagic septicemia freedom: A decision science approach. *Preventive Veterinary Medicine, 94*(1–2), 140–153. doi:10.1016/j.prevetmed.2009.11.021.

Gutierrez, B. T., Plant, N. G., & Thieler, E. R. (2011). A Bayesian network to predict vulnerability to sea-level rise: Data report. *U.S. Geological Survey Data Series 601*.

Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(5), 1366–1373. doi:10.1037/a0016567.

Heuze, Y., & Braga, J. (2008). Application of non-adult Bayesian dental age assessment methods to skeletal remains: The Spitalfields collection. *Journal of Archaeological Science, 35*(2), 368–375. doi:10.1016/j.jas.2007.04.003.

Hoffman, B., & Schraw, G. (2010). Conceptions of efficiency: Applications in learning and problem solving. *Educational Psychologist, 45*(1), 1–14. doi:10.1080/00461520903213618.

Hood, G. M., Barry, S. C., & Martin, P. A. J. (2009). Alternative methods for computing the sensitivity of complex surveillance systems. *Risk Analysis, 29*(12), 1686–1698. doi:10.1111/j.1539-6924.2009.01323.x.

Hossain, M., & Muromachi, Y. (2012). A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. *Accident Analysis and Prevention, 45*, 373–381. doi:10.1016/j.aap. 2011.08.004.

Hsu, C. I., Shih, M. L., Huang, B. W., Lin, B. Y., & Lin, C. N. (2009). Predicting tourism loyalty using an integrated Bayesian network mechanism. *Expert Systems with Applications, 36*(9), 11760–11763. doi:10.1016/j.eswa.2009.04.010.

Inman, D., Blind, M., Ribarova, I., Krause, A., Roosenschoon, O., Kassahun, A., et al. (2011). Perceived effectiveness of environmental decision support systems in participatory planning: Evidence from small groups of end-users. *Environmental Modelling and Software, 26*(3), 302–309. doi:10.1016/j.envsoft.2010.08.005.

Jalen, L., & Mamon, R. (2009). Valuation of contingent claims with mortality and interest rate risks. *Mathematical and Computer Modelling, 49*(9–10), 1893–1904. doi:10.1016/j.mcm.2008.10.014.

Joseph, S. A., Adams, B. J., & McCabe, B. (2010). Methodology for Bayesian belief network development to facilitate compliance with water quality regulations. *Journal of Infrastructure Systems, 16*(1), 58–65. doi:10.1061/(asce)1076-0342(2010) 16:1(58).

Joyce, J. (2008). *"Bayes' Theorem", The Stanford encyclopedia of philosophy* (Fall 2008 Edition), Edward N. Zalta (Ed.), http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/

Kasess, C. H., Stephan, K. E., Weissenbacher, A., Pezawas, L., Moser, E., & Windischberger, C. (2010). Multi-subject analyses with dynamic causal modeling. *NeuroImage, 49*(4), 3065–3074. doi:10.1016/j.neuroimage.2009.11.037.

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32*(4), 995–1003. doi:10.1016/j.eswa.2006.02.016.

Kisioglu, P., & Topcu, Y. I. (2011). Applying Bayesian belief network approach to customer churn analysis: A case study on the telecom industry of Turkey. *Expert Systems with Applications, 38*(6), 7151–7157. doi:10.1016/j.eswa.2010.12.045.

Kleiter, G. D. (2007). Implications of natural sampling in base-rate tasks. *The Behavioral and Brain Sciences, 30*(3), 270–271. doi:10.1017/s0140525x07001793 [Editorial Material].

Koelle, D., Pfautz, J., Farry, M., Cox, Z., Catto, G., & Campolongo, J. (2006). *Applications of Bayesian belief networks in social network analysis*. Charles River Analytics Inc. 22nd Annual Conference on Uncertainty in Artificial Intelligence: UAI '06, Cambridge, MA.

Kragt, M. E., Newham, L. T. H., Bennett, J., & Jakeman, A. J. (2011). An integrated approach to linking economic valuation and catchment modelling. *Environmental Modelling and Software, 26*(1), 92–102. doi:10.1016/j.envsoft.2010.04.002.

Kusumastuti, D., Hannes, E., Depaire, B., Vanhoof, K., Janssens, D., Wets, G., et al. (2011). An interactive computer-based interface to support the discovery of individuals' mental representations and preferences in decisions problems: An application to travel. *Computers in Human Behavior, 27*(2), 997–1011. doi:10.1016/j.chb.2010.12.004.

Laming, D. (2007). Ordinary people do not ignore base rates. *The Behavioral and Brain Sciences, 30*(3), 272–274. doi:10.1017/s0140525x0700181 [Editorial Material].

Lee, K. J., & Chang, W. (2009). Bayesian belief network for box-office performance: A case study on Korean movies. *Expert Systems with Applications, 36*(1), 280–291. doi:10.1016/j.eswa.2007.09.042.

Lehmann, E. L. (1995). Neyman's statistical philosophy. *Probability and Mathematical Statistics, 15*, 29–36.

Levontin, P., Kulmala, S., Haapasaari, P., & Kuikka, S. (2011). Integration of biological, economic, and sociological knowledge by Bayesian belief networks: The interdisciplinary evaluation of potential management plans for Baltic salmon. *ICES Journal of Marine Science, 68*(3), 632–638. doi:10.1093/icesjms/fsr004.

Linders, B. (2009). Building process improvement business cases using Bayesian belief networks and Monte Carlo simulation. *Software Engineering Institute, CMU/SEI-2009-TN-017*.

Lindgaard, G., Pyper, C., Frize, M., & Walker, R. (2009). Does Bayes have it? Decision support systems in diagnostic medicine. *International Journal of Industrial Ergonomics, 39*(3), 524–532. doi:10.1016/j.ergon.2008.10.011.

Little, T. T., Miller, G., & Guilmette, R. (2003). Internal dosimetry of uranium isotopes using Bayesian inference methods. *Radiation Protection Dosimetry, 105*(1–4), 413–416.

Liu, Q., Otter, T., & Allenby, G. M. (2007). Investigating endogeneity bias in marketing. *Marketing Science, 26*(5), 642–650. doi:10.1287/mksc.1060.0256.

Louvieris, P., Gregoriades, A., & Garn, W. (2010). Assessing critical success factors for military decision support. *Expert Systems with Applications, 37*(12), 8229–8241. doi:10.1016/j.eswa.2010.05.062.

Loyd, K. A. T., & DeVore, J. L. (2010). An evaluation of Feral cat management options using a decision analysis network. *Ecology and Society, 15*(4): 10. http://www.ecologyandsociety.org/vol15/iss4/art10/.

Luu, V. T., Kim, S. Y., Nguyen, V. T., & Ogunlana, S. O. (2009). Quantifying schedule risk in construction projects using Bayesian belief networks. *International Journal of Project Management, 27*(1), 39–50. doi:10.1016/j.ijproman.2008.03.003.

Ma, L., Arentze, T., Borgers, A., & Timmermans, H. (2007). Modelling land-use decisions under conditions of uncertainty. *Computers Environment and Urban Systems, 31*(4), 461–476. doi:10.1016/j.compenvurbsys.2007.02.002.

Ma, J. M., Kockelman, K. M., & Damien, P. (2008). A multivariate poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis and Prevention, 40*(3), 964–975. doi:10.1016/j.aap. 2007.11.002.

Macauley, M., & Laxminarayan, R. (2010). The value of information: 'Methodological frontiers and new applications for realizing social benefit' workshop. *Space Policy, 26*(4), 249–251. doi:10.1016/j.spacepol.2010.08.007.

Mannor, S., & Shinikin, N. (2008). Regret minimization in repeated matrix games with variable stage duration. *Games and Economic Behavior, 63*(1), 227–258. doi:10.1016/j.geb.2007.07.006.

Marhavilas, P. K., Koulouriotis, D., & Gemeni, V. (2011). Risk analysis and assessment methodologies in the work sites: On a review, classification and comparative study of the scientific literature of the period 2000–2009. *Journal of Loss Prevention in the Process Industries, 24*(5), 477–523. doi:10.1016/j.jlp. 2011.03.004 [Review].

Martin, J. E., Rivas, T., Matias, J. M., Taboada, J., & Arguelles, A. (2009). A Bayesian network analysis of workplace accidents caused by falls from a height. *Safety Science, 47*(2), 206–214. doi:10.1016/j.ssci.2008.03.004.

Martinez, E. Z., Ruffino-Netto, A., Achcar, J. A., & Aragon, D. C. (2008). Bayesian model for the risk of tuberculosis infection for studies with individuals lost to follow-up. *Revista De Saude Publica, 42*(6), 999–1004.

Mast, B. D. (2010). Measuring neighborhood quality with survey data: A Bayesian approach. *Cityscape: A Journal of Policy Development and Research, 12*(3). 123–143.

Mast, B. D. (2012). Comparing public housing and housing voucher tenants with Bayesian propensity scores. *Cityscape, 14*(1), 55–72.

Mazzocco, T., & Hussain, A. (2012). Novel logistic regression models to aid the diagnosis of dementia. *Expert Systems with Applications, 39*(3), 3356–3361. doi:10.1016/j.eswa.2011.09.023.

McCabe, B., Loughlin, C., Munteanu, R., Tucker, S., & Lam, A. (2008). Individual safety and health outcomes in the construction industry. *Canadian Journal of Civil Engineering, 35*(12), 1455–1467. doi:10.1139/l08-091.

McCarron, C. E., Pullenayegum, E. M., Marshall, D. A., Goeree, R., & Tarride, J. E. (2009). Handling uncertainty in economic evaluations of patient level data: A review of the use of Bayesian methods to inform health technology assessments. *International Journal of Technology Assessment in Health Care, 25*(4), 546–554. doi:10.1017/s0266462309990316 [Review].

McCloskey, J. T., Lilieholm, R. J., & Cronan, C. (2011). Using Bayesian belief networks to identify potential compatibilities and conflicts between development and landscape conservation. *Landscape and Urban Planning, 101*(2), 190–203. doi:10.1016/j.landurbplan.2011.02.011.

McGrayne, S. B. (2011). *The theory that would not die: How Bayes' rule cracked the Enigma code, hunted down Russian submarines, and emerged triumphant from two centuries of controversy*. New Haven: Yale University Press.

McKaughan, D. C., Heath, Z., & McClain, J. T. (2011). *Using a text analysis and categorization tool to generate Bayesian belief networks for use in cognitive social simulation from a document corpus.* Paper presented at the Proceedings of the 2011 Military Modeling & Simulation Symposium, Boston, MA

Michaeli, R., & Simon, L. (2008). An illustration of Bayes' theorem and its use as a decision-making aid for competitive intelligence and marketing analysts. *European Journal of Marketing, 42*(7–8), 804–813. doi:10.1108/03090560810877169.

Milgrom, J., Mendelsohn, J., & Gemmill, A. W. (2011). Does postnatal depression screening work? Throwing out the bathwater, keeping the baby. *Journal of Affective Disorders, 132*(3), 301–310. doi:10.1016/j.jad.2010.09.031 [Review].

Mohaghegh, Z., Kazemi, R., & Mosleh, A. (2009). Incorporating organizational factors into probabilistic risk assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. *Reliability Engineering and System Safety, 94*(5), 1000–1018. doi:10.1016/j.ress.2008.11.006.

Murray-Tuite, P. M., & Fei, X. A. (2010). A methodology for assessing transportation network terrorism risk with attacker and defender interactions. *Computer-Aided Civil and Infrastructure Engineering, 25*(6), 396–410. doi:10.1111/j.1467-8667.2010.00655.x.

Murungweni, C., van Wijk, M. T., Andersson, J. A., Smaling, E. M. A., & Giller, K. E. (2011). Application of fuzzy cognitive mapping in livelihood vulnerability analysis. *Ecology and Society, 16*(4). doi:10.5751/es-04393-160408.

Nasoz, F., Lisetti, C. L., & Vasilakos, A. V. (2010). Affectively intelligent and adaptive car interfaces. *Information Sciences, 180*(20), 3817–3836. doi:10.1016/j.ins.2010.06.034.

Neil, M., Fenton, N., Forey, S., & Harris, R. (2001). Using Bayesian belief networks to predict the reliability of military vehicles. *Computing & Control Engineering Journal, 12*(1), 11–20. doi:10.1049/ccej:20010103.

Neil, M., Hager, D., & Andersen, L. B. (2009). Modeling operational risk in financial institutions using hybrid dynamic Bayesian networks. *Journal of Operational Risk, 4*(1), 3–33.

Newton, A. C., Marshall, E., Schreckenberg, K., Golicher, D., Te Velde, D. W., Edouard, F., et al. (2006). Use of a Bayesian belief network to predict the impacts of commercializing non-timber forest products on livelihoods. *Ecology and Society*, 11(2), 24. http://www.ecologyandsociety.org/vol11/iss2/art24/.

Newton, A. C., Cayuela, L., Echeverria, C., Armesto, J. J., Del Castillo, R. F., Golicher, D., et al. (2009). Toward integrated analysis of human impacts on forest biodiversity: Lessons from Latin America. *Ecology and Society, 14*(2): 2. http://www.ecologyandsociety.org/vol14/iss2/art2/.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y. J., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems, 50*(3), 559–569. doi:10.1016/j.dss.2010.08.006.

Norrington, L., Quigley, J. L., Russell, A. H., & Van Der Meer, R. B. (2008). Modelling the reliability of search and rescue operations with Bayesian belief networks. *Reliability Engineering and System Safety, 93*(7), 940–949. doi:10.1016/j.ress.2007.03.006.

Oakil, A. M., Ettema, D., Arentze, T., & Timmermans, H. (2011). Longitudinal model of longer-term mobility decisions: Framework and first empirical tests. *Journal of Urban Planning and Development-ASCE, 137*(3), 220–229. doi:10.1061/(asce)up. 1943-5444.0000066.

Oaksford, M., & Chater, N. (2009). Precis of Bayesian rationality: The probabilistic approach to human reasoning. *The Behavioral and Brain Sciences, 32*(1), 69. doi:10.1017/s0140525x09000284 [Review].

Padilla, C., Lalloue, B., Zmirou-Navier, D., & Severine, D. (2011). Infant mortality, deprivation and proximity to polluting industrial facilities – a small-scale spatial analysis with census data (Lille metropolitan area, France). *Environment Risques & Sante, 10*(3), 216–221. doi:10.1684/ers.2011.0455.

Pearl, J. (1985). A model of self-activated memory for evidential reasoning. In *7th Conference of the Cognitive Science Society, University of California, Irvine, August 1985* (pp. 329–334)

Peterson, L. T., Ford, E. W., Eberhardt, J., Huerta, T. R., & Menachemi, N. (2011). Assessing differences between physicians' realized and anticipated gains from electronic health record adoption. *Journal of Medical Systems, 35*(2), 151–161. doi:10.1007/s10916-009-9352-z.

Raiffa, H. (2012). *Howard Raiffa on decision trees*. http://www.hbs.edu/centennial/im/inquiry/sections/3/b/page15.html. Accessed 26 June 2012.

Ranta, J., Matjushin, D., Virtanen, T., Kuusi, M., Viljugrein, H., Hofshagen, M., et al. (2011). Bayesian temporal source attribution of foodborne zoonoses: Campylobacter in Finland and Norway. *Risk Analysis, 31*(7), 1156–1171. doi:10.1111/j.1539-6924.2010.01558.x.

Replogle, W. H., Johnson, W. D., & Hoover, K. W. (2009). Using evidence to determine diagnostic test efficacy. *Worldviews on Evidence-Based Nursing, 6*(2), 87–92.

Rupp, J. D., Flannagan, C. A. C., Hoff, C. N., & Cunningham, R. M. (2010). Effects of osteoporosis on AIS 3 + injury risk in motor-vehicle crashes. *Accident Analysis and Prevention, 42*(6), 2140–2143. doi:10.1016/j.aap. 2010.07.005.

Saravanan, V. S. (2010). Negotiating participatory irrigation management in the Indian Himalayas. *Agricultural Water Management, 97*(5), 651–658. doi:10.1016/j.agwat.2009.12.003.

Scarlata, S., Pedone, C., Conte, M. E., & Incalzi, R. A. (2009). Accuracy of spirometry in diagnosing pulmonary restriction in elderly people. *Journal of the American Geriatrics Society, 57*(11), 2107–2111. doi:10.1111/j.1532-5415.2009.02525.x.

Schneider, I. (2005). Abraham De Moivre, the doctrine of chances (1718, 1738, 1756). In I. Grattan-Guinness (Ed.), *Landmark writings in western mathematics 1640–1940*. Amsterdam: Elsevier.

Schultz, G., Thurgood, D., Olsen, A., & Reese, C. S. (2011). *Analyzing raised median safety impacts using Bayesian methods*. Presented at the 90th meeting of the Transportation Research Board, Washington, DC.

Sebastiani, P., Ramoni, M., Cohen, P., Warwick, J., & Davis, J. (1999). Discovering dynamics using Bayesian clustering. In D. Hand, J. Kok, & M. Berthold (Eds.), *Advances in intelligent data analysis* (Lecture notes in computer science, Vol. 1642, pp. 199–209). Berlin/Heidelberg: Springer.

Sevcikova, H., Raftery, A. E., & Waddell, P. A. (2007). Assessing uncertainty in urban simulations using Bayesian melding. *Transportation Research Part B-Methodological, 41*(6), 652–669. doi:10.1016/j.trb.2006.11.001.

Small, M. J. (2008). Methods for assessing uncertainty in fundamental assumptions and associated models for cancer risk assessment. *Risk Analysis, 28*(5), 1289–1307. doi:10.1111/j.1539-6924.2008.01134.x.

Smid, J. H., Swart, A. N., Havelaar, A. H., & Pielaat, A. (2011). A practical framework for the construction of a biotracing model: Application to Salmonella in the pork slaughter chain. *Risk Analysis, 31*(9), 1434–1450. doi:10.1111/j.1539-6924.2011.01591.x.

Smith, C., Felderhof, L., & Bosch, O. J. H. (2007). Adaptive management: Making it happen through participatory systems analysis. *Systems Research and Behavioral Science, 24*(6), 567–587. doi:10.1002/sres.835.

Social Sciences Citation Index®, accessed via Web of Science® (2012). THOMSON REUTERS. Accessed June 1, 2012. http://thomsonreuters.com/products_services/science/science_products/a-z/social_sciences_citation_index/

Staker, R. J. (1999). Use of Bayesian belief networks in the analysis of information system network risk. *Information, Decision and Control, 1999. IDC 99. Proceedings. 1999,* (pp. 145–150). doi:10.1109/idc.1999.754143.

Starr, C., & Shi, P. (2004). An introduction to Bayesian belief networks and their applications to land operations. *Network, 391*(6670). doi:10.1038/36103.

Stelzenmuller, V., Schulze, T., Fock, H. O., & Berkenhagen, J. (2011). Integrated modelling tools to support risk-based decision-making in marine spatial management. *Marine Ecology Progress Series, 441*, 197–212. doi:10.3354/meps09354.

Swiler, L. P. (2006). Bayesian methods in engineering design problems. *Optimization, SAND 2005-3294*.

Tang, D. W., Yang, J. B., Chin, K. S., Wong, Z. S. Y., & Liu, X. B. (2011). A methodology to generate a belief rule base for customer perception risk analysis in new product development. *Expert Systems with Applications, 38*(5), 5373–5383. doi:10.1016/j.eswa.2010.10.018.

Thodi, P. N., Khan, F. I., & Haddara, M. R. (2010). The development of posterior probability models in risk-based integrity modeling. *Risk Analysis, 30*(3), 400–420. doi:10.1111/j.1539-6924.2009.01352.x.

Vans, O. (1998). A belief network approach to optimization and parameter estimation: Application to resource and environmental management. *Artificial Intelligence, 101*(1–2), 135–163. doi:10.1016/s0004-3702(98)00010-1.

Verma, M. (2011). Railroad transportation of dangerous goods: A conditional exposure approach to minimize transport risk. *Transportation Research Part C-Emerging Technologies, 19*(5), 790–802. doi:10.1016/j.trc.2010.07.003.

Walshe, T., & Burgman, M. (2010). A framework for assessing and managing risks posed by emerging diseases. *Risk Analysis, 30*(2), 236–249. doi:10.1111/j.1539-6924.2009.01305.x.

Wollert, R. (2007). Poor diagnostic reliability, the null-Bayes logic model, and their implications for sexually violent predator evaluations. *Psychology, Public Policy, and Law, 13*(3), 167–203. doi:10.1037/1076-8971.13.3.167.

Wong, W., Liu, W., & Bennamoun, M. (2009). A probabilistic framework for automatic term recognition. *Intelligent Data Analysis, 13*(4), 499–539. doi:10.3233/ida-2009-0379.

Zellner, A. (2002). Information processing and Bayesian analysis. *Journal of Econometrics, 107* (1–2), 41–50. doi:10.1016/s0304-4076(01)00112-9.

Zimmerman, L. A., Sestokas, J. M., Burns, C. A., Grover, J., Topaz, D., & Bell, J. (2010). *Improving attrition rates in the M1A1/M1A2 Master Gunner course* (Final report prepared under Subcontract ARI-ARA-07-001, Task order 14524 (DRC); Contract # W74V8H-04-D-0048 for U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Knox, KY). Fairborn, OH: Klein Associates Division of Applied Research Associates.

Zorrilla, P., Carmona, G., De la Hera, A., Varela-Ortega, C., Martinez-Santos, P., Bromley, J., et al. (2010). Evaluation of Bayesian networks in participatory water resources management, Upper Guadiana Basin, Spain. *Ecology and Society, 15*(3):12. http://www.ecologyandsociety.org/vol15/iss3/art12/.

Chapter 3
Statistical Properties of Bayes' Theorem

3.1 Introduction to Statistical Terminology

This chapter provides a review of the basis statistical properties associated with modeling BBN. I begin with a simple two Bayes' theorem proofs to show the conditional and unconditional relationships between them in revising the priors to formulate *a-posterior or revised* a-priori probabilities. This revision process begins with the priors, is filtered through the likelihood, joint, and marginal probabilities and finishes with the a-posterior probabilities.

3.2 Bayes' Theorem Proof

I present a two and three event proof of Bayes' theorem.[1] The two event proof validates the BBN models in Chaps. 5, 6, 7, 8, 9, 10, 11, 12 and 13 and the three event proof, the BBN model in Chap. 14.

3.2.1 A Bayes' Theorem Proof with Two Events, A and B

Given, $(A \cap B) = (A \cap B)$, it follows from the chain rule and conditional probability that:

- Step 1: $P(B \cap A) = P(B)\,P(A|B)$ and
- Step 2: $P(A \cap B) = P(A)\,P(B|A)$, then

[1] These proofs only represent one path across a BBN. For example, if a BBN consists of an Event B with two sub-elements and an Event A with two sub-elements, then there are 2×4 or 4 total paths. If the Event A has three sub-elements, then there would be 2×3 or 6 paths, etc. For example, this proof traces the event path $B \rightarrow$ A. Other paths for a $2 \times 2 \times 2$ BBN include B $\rightarrow \bar{A}$ or $\bar{B} \rightarrow \bar{A}$, etc.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_3,

- Step 3: $P(B)\,P(A|B) = P(A)\,P(B|A)$, where
- Step 4: $P(B|A) = \frac{P(B)P(A|B)}{P(A)} = \frac{P(A \cap B)}{P(A)}$, Bayes' theorem, *qed*.

Bolstad (2007) suggest a general form as:

$P(B_i|A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\Sigma_{j=1}^{n} P(A|B_j)P(B_j)}$, where P(A) and P(B) ≥ 0

and P(B_i) consists of mutually exclusive (disjoint) events within the universe, S.[2]

The elegance of this simple proof is that it allows one to transition from the truth of an Event B given the evidence contained in Event A or from the truth contained in Event A given the truth of Event B. Its utility in learning is that the truth contained in the *a posteriori* P(B|A) becomes the a-priori truth for the next iteration of a chained BBN.

3.2.2 A Step-by-Step Explanation of the Two Event Bayes' Theorem Proof

I begin with setting $(A \cap B) = (A \cap B)$ and using the chain (product) rule of probability.

Step 1: $P(B \cap A) = P(B|A)\,P(A)$, using the chain rule.

Here, the joint probability of events B and A, $P(B \cap A)$, is equal to the (conditional) probability of an Event B given the simultaneously occurrence of Event A, $P(B|A)$ times the probability of an Event A, P(A).

Step 2: $P(A \cap B) = P(A|B)\,P(B)$, using the chain rule.

Here, the joint probability of events A and B, $P(A \cap B)$, is equal to the (conditional) probability of an Event A given the simultaneously occurrence of Event B, $P(A|B)$, times the probability of an Event B, P(B).

Given $P(B \cap A) = P(A \cap B)$

Step 3: P(B|A) P(A) = P(A|B) P(B) where

$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$, and rearranging,

Step 4: $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$, qed.

Now, the conditional probability of Event B, given the probability of a given Event A, P(B|A) is equal to the conditional probability of Event A given Event B P(A|B) times the probability of Event B, P(B), divided by the probability of Event A, P(A), which is Bayes' theorem.

[2] Note the law of total probability allows us to reform P(A) into $\sum_{j=1}^{n} P(A \cap B_j)$ or $P(A) = \sum_{j=1}^{n} P(A|B_j) \times P(B_j)$ (Bolstad 2007).

3.2.3 A Bayes' Theorem Proof with Three Events, A, B, & C

Given, $(A_i \cap B \cap C) = (C \cap B \cap A_i)$, it follows from the chain rule and conditional probability that:

- Step 1: $P(A_i \cap B \cap C) = P(A_i|B \cap C)\ P(B|C)\ P(C),$ and
- Step 2: $P(C \cap B \cap A_i) = P(C|B \cap A_i)\ P(B|A_i)\ P(A_i),$ and
- Step 3: $P(A_i|B \cap C)\ P(B|C)\ P(C) = P(C|B \cap A_i)\ P(B|A_i)\ P(A_i).$
- Step 4: $P(A_i|B \cap C) = \frac{P(C|B\ \cap\ A_i)\ P(B|A_i)P(A_i)}{P(B|C)\ P(C)} = \frac{P(C|B\ \cap\ A_i)\ P(B\ \cap\ A_i)}{(B\ \cap\ C)} = \frac{(A_i\ \cap\ B\ \cap\ C)}{(B\ \cap\ C)}$, Bayes' theorem, *qed*.[3]

To be able to use this across a BBN, we would to invoke the Law of Total Probability for conditional events.

This done as follows:

From $\frac{P(C|B\ \cap\ A_i)\ P(B\ \cap\ A_i)}{(B\ \cap\ C)}$, we remove the nuisance parameter B again by using the Chain Rule of probability. We begin by reversing $P(B \cap A_i)$ and $(B \cap C)$, $\frac{P(C|B\ \cap\ A_i)\ P(A_i \cap\ B)}{(C\ \cap\ B)}$, and applying the Chain Rule again and eliminate P(B): $\frac{P(C|B\ \cap\ A_i)\ P(|A|B)P(B)}{P(C|B)P(B)}$. Now we apply the Law of Total Probability for conditional events as such:

$$P(C|B) = \sum_i P(C|B \cap A_i)P(A_i|B).$$

Expanding this we have for i = A and $\bar{A}$, $P(C|B)$ = P(C|B ∩ A)P(A|B)+ {P(C |B ∩ $\bar{\text{A}}$)P($\bar{\text{A}}$ |B).[4]

Now, we have: $P(A|B \cap C) = \frac{P(C|B\ \cap\ A_i)\ P(|A|B)}{\text{P(C|B } \cap\ A\text{) P(A|B)+P(C|B } \cap \bar{A})P(\bar{A}\ |B)}$.

3.2.4 Independence and Conditional Independence Evaluation

3.2.4.1 Independence

Here is where we can test for independence. If P(A|B) = P(A) or P(B|A) = P(B), then the two events are independent. If they are independent, then there is not a cause and effect relationship between the events; i.e., Event A does not cause Event B. If this is true, there is no utility in using Bayes' theorem as a predictive tool.

[3] Using this proof, using the Law of Total Probability, we are only interested in solving across two BBN paths: A→B→C and $\bar{\text{A}}$→B→C.

[4] Gregory (2005) offers this as a "Usual Form" of Bayes' Theorem.

3.2.4.2 Conditional Independence

Here, I will show that the conditional independence property of the relationship below is true. This relationship is one in which I condition Event A and B on an Event C and provide a proof. I am determining if the joint probability of Events A and B conditioned or given the marginal probability of Event C is equal to the joint probability of Events A, B, and C divided by the marginal probability of Event C.

- Step 1: $P(A \cap B|C) = \dfrac{P(A \cap B \cap C)}{P(C)}$. Rearranging:
- Step 2: $P(A \cap B \cap C) = P(A \cap B|C)\,P(C)$, from the chain rule:
- Step 3: $\mathrm{P(A \cap B \cap C) = P(A)\,P(B|A)\,P(C|A \cap B)}$. Given
- Step 4: $\mathrm{P(A \cap B \cap C) = P(A \cap B \cap C)}$, then
- Step 5: $\mathrm{P(A \cap B|C)\ P(C) = P(A)\,P(B|A)\,P(C|A \cap B)}$, and
- Step 6: $P(A \cap B|C) = \dfrac{\mathrm{P(A)P(B|A)\,P(C|A \cap B)}}{P(C)}$, and from the chain rule

  $P(A \cap B) = \mathrm{P(A)\,P(B|C)}$, then

- Step 7: $P(A \cap B|C) = \dfrac{\mathrm{P(A \cap B)\,P(C|A \cap B)}}{P(C)}$, and
- Step 8: $P(A \cap B|C) = \dfrac{\mathrm{P(A \cap B \cap C)}}{P(C)}$, qed.

3.3 Statistical Definitions

3.3.1 Axioms of Probability

The following are three basic axioms of probability that underpin the concepts of Bayes' theorem and BBN: (1) chances are always at least zero (never negative), $\mathrm{P(A)} \geq 0$; (2) the chance that something happens in a universe is always 100%, $\mathrm{P}(\cup) = 1$; and (3) if two events cannot both occur at the same time (if they are disjoint or mutually exclusive), the chance that either one occurs is the sum of the chances that each occurs, $\mathrm{P(A \cup B) = P(A) + P(B)}$. For non-mutually events this is $\mathrm{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$. We can derive other mathematical facts about probability from these three axioms.

3.3.2 Bayes' Theorem

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

where the conditional probability of Event B, given the probability of a given Event A, P(B|A) is equal to the conditional probability of Event A given Event B P(A|B) times the probability of Event B, P(B), divided by the probability of Event A, P(A), which is Bayes' theorem.

3.3.3 Combinations and Permutations

Combinations. The number of combinations of n things taken k at a time is the number of ways of picking a subset of k of the n things, without replacement, and without regard to the order in which we select the elements of the subset. The number of such combinations is ${}_nC_k = n!/(k!\ (n-k)$, where k! (pronounced "k Factorial") is $k \times (k-1) \times (k-2) \times \ldots \times 1$. The numbers ${}_nC_k$ are also called the *Binomial coefficients*. From a set that has n elements one can form a total of 2^n subsets of all sizes. For example, from the set {a, b, c}, which has 3 elements, one can form the $2^3 = 8$ subsets { }, {a}, {b}, {c}, {a, b}, {a, c}, {b, c}, {a, b, c}. Because the number of subsets with k elements one can form from a set with n elements is ${}_nC_k$, and the total number of subsets of a set is the sum of the numbers of possible subsets of each size, it follows that ${}_nC_0 + {}_nC_1 + {}_nC_2 + \ldots + {}_nC_n = 2^n$. **Permutations**. A permutation of a set is an arrangement of the elements of the set in some order. If the set has n things in it, there are N! different orderings of its elements. For the first element in an ordering, there are n possible choices, for the second, there remain n − 1 possible choices, for the third, there are n − 2, etc., and for the nth element of the ordering, there is a single choice remaining. By the fundamental rule of counting, the total number of sequences is thus n × (n − 1) × (n − 2) × ... × 1. Similarly, the number of orderings of length k one can form from n ≥ k things is n × (n − 1) × (n − 2) × ... × (n − k + 1) = n!/(n − k)!. This is denoted ${}_nP_k$, the number of permutations of n things taken k at a time.

3.3.4 Conditional and Unconditional Probability

Conditional Probability. Suppose there is some interest in the probability that some Event A occurs, and we learn that the Event B occurred. How should we update the probability of A to reflect this new knowledge? This is what the conditional probability does: it says how the additional knowledge that B occurred should affect the probability that A occurred quantitatively. For example, suppose that A and B are mutually exclusive. Then if B occurred, A did not, so the conditional probability that A occurred given that B occurred is zero. At the other extreme, suppose that B is a subset of A, so that A must occur whenever B does. Then if we learn that B occurred, A must have occurred too, so the conditional probability that A occurred given that B occurred is 100%. For in-between cases, where A and B intersect, but B is not a subset of A, the conditional probability of A given B is a number between zero and one. Basically, one "restricts" the outcome space S to consider only the part of S that is in B, because we know that B occurred. For A to have happened given that B happened requires that AB happened, so we are interested in the Event AB. To have a legitimate probability requires that P (S) = 100%, so if we are restricting the outcome space to B, we need to divide by the probability of B to make the probability of this new S be 100%. On this scale, the probability that AB happened is P(AB) / P(B). This is the definition of the conditional probability of A given B, provided P(B) is not zero (division by zero is undefined). Note, that the special cases AB = { } (A and B are mutually exclusive) and AB = B (B is a subset of A) agree with our intuition as described at the top of this paragraph. Conditional probabilities satisfy the axioms of probability, just as ordinary probabilities do. **Unconditional Probability**. This is the probability of B given A, which is that part of B that is also in A. When we multiply it by a scale factor 1/P(A) it becomes a conditional probability (Bolstad (2007)).

3.3.5 Counting, Countable and Uncountable Set

Counting. To count a set of things is to put it in one to one correspondence with a consecutive subset of the positive integers, starting with 1. **Countable Set**. A set is countable if we can place its elements in one-to-one correspondence with a subset of the integers. For example, the sets {0, 1, 7, −3}, {red, green, blue}, {...,−2, −1, 0, 1, 2, ...}, {straight, curly}, and the set of all fractions, are countable. If a set is not countable, it is uncountable. **Uncountable Set**. A set is uncountable if it is not countable. The set of all real numbers is uncountable.

3.3.6 Complement and Complement Rule

Complement. The complement of a subset of a given set is the collection of all elements of the set that are not elements of the subset. **Complement Rule**. This states that the probability of the complement of an event is 100% minus the probability of the event: $P(A^c) = 100\% - P(A)$.

3.3.7 Disjoint or Mutually Exclusive Events/Sets

Events. Two events are disjoint or mutually exclusive if the occurrence of one is incompatible with the occurrence of the other; that is, if they cannot both happen at once (if they have no outcome in common). Equivalently, two events are disjoint if their intersection is the empty set. **Sets**. Two sets are disjoint or mutually exclusive if they have no element in common. Equivalently, two sets are disjoint if their intersection is the empty set.

3.3.8 Event

An event is a subset of outcome space. An event determined by a random variable has the form A = (X is in A). When we observe the random variable X, that determines if A will occur: if the value of X is in A, A occurs; if not, A does not occur.

3.3.9 Factorial

For an integer k that is greater than or equal to 1, $k!$ (pronounced "k factorial") is $k \times (k - 1) \times (k - 2) \times \ldots \times 1$. By convention, $0! = 1$. There are $k!$ ways of ordering k distinct objects. For example, 9! is the number of batting orders of 9 baseball players, and 52! is the number of different ways one can order a standard deck of playing cards.

3.3.10 Intersection and Union (of Sets)

Intersection The intersection of two or more sets is the set of elements that all the sets have in common; the elements contained in every one of the sets. The intersection of the events A and B is written, "$A \cap B$". **Union**. The union of two or more sets is the set of objects contained by at least one of the sets. I denote the union of the events A and B as, "$A \cup B$."

3.3.11 Joint and Marginal Probability Distribution

Joint If $X_1, X_2, \ldots, X_k$ are random variables defined for the same experiment, their joint probability distribution gives the probability of events determined by the collection of random variables: for any collection of sets of numbers $\{A_1, \ldots, A_k\}$, the joint probability distribution determines P((X_1 is in A_1) and (X_2 is in A_2) and … and (X_k is in A_k). **Marginal**. The marginal probability distribution of a random variable that has a joint probability distribution with some other random variables is the probability distribution of that random variable without regard for the values that the other random variables take. One can find the marginal distribution of a discrete random variable X_1 that has a joint distribution with other discrete random variables from the joint distribution by summing over all possible values of the other variables. **Marginalization.**[5] A technique in BBN when dealing with an unknown parameter and conditional probabilities using the Total Law of Probability. Using a three event example, we have:

$$P(A_i|B \cap C) = \frac{P(A_i|C)P(B|A_i \cap C)}{P(B|C)}$$

Here, we can update our belief in hypothesis A_i given the additional evidence B and the background information C. The left-hand term, $P(A_i|B \cap C)$ is the posterior probability, or the probability of A_i after an analyst considers the effect of B given C. The term $P(A_i|C)$ is the prior probability of A_i given C alone. The term $P(B|A_i \cap C)$ is the likelihood and gives the probability of the evidence assuming the hypothesis A and the background information C is true. Finally, the last term P(B|C) is the expectedness, or how expected the evidence is given only C. It is independent of A_i and an analyst can regard it as a marginalizing or scaling factor.

We can rewrite this as $P(B|C) = \sum_i P(B|A_i \cap C)P(A_i|C)$ where i denotes a specific hypothesis A_i, and the summation is taken over a set of hypotheses which are mutually exclusive and exhaustive (their prior probabilities sum to 1). Hebert et al. (2007) also suggests it is important to note that all of these probabilities are conditional in that they specify the degree of belief in some proposition or propositions based on an initial assumption that some other propositions are true. So, this prior determination of the probability of these previous propositions underpins the theory.

3.3.12 Mean, Arithmetic Mean

The sum of a list of numbers, divided by the number of numbers.

3.3.13 Outcome Space

The outcome (outcome space) is the set of all possible outcomes of a given random experiment and I will denote this by the capital letter S.

[5] I derived this definition from Hebert et al. (2007).

3.3.14 Parameter

A numerical property of a population, such as its mean.

3.3.15 Partition

A *partition* of an Event A is a collection of events $\{A_1, A_2, A_3, \ldots\}$ such that the events in the collection are disjoint and their union is A. That is,

$$A_j A_k = \{\} \text{ unless } j = k, \text{ and } A = A_1 \cup A_2 \cup A_3 \cup \ldots.$$

If the Event A is not specified, we assume it to be the entire outcome **S**.

3.3.16 Population

This is a collection of studied units. Units can be people, places, objects, epochs, drugs, procedures, or many other things. Much of statistics is concerned with estimating numerical properties (parameters) of an entire population from a random sample of units from the population.

3.3.17 Prior and Posterior Probabilities

See Bayes' theorem.

3.3.18 Probability and Probability Sample

Probability The probability (chance) of an event is a real number between zero and 100%. The meaning (interpretation) of probability is the subject of theories of probability, which differ in their interpretations. However, any rule for assigning probabilities to events has to satisfy the axioms of probability. **Probability Sample**. A draw from a population using a random mechanism so that every element of the population has a known chance of ending up in the sample.

3.3.19 Product (Chain Rule)

The Chain Rule for Probability for non-zero events allows for the expansion of members of a set of random variables from a joint distribution across any BBN using only conditional probabilities. For example, if we have a collection of events $A_1 \ldots A_n$, we can expand this as:

$$P(A_n, \ldots, A_1) = P(A_n | \; A_{n-1}, \ldots, A_1) P(A_{n-1}, \ldots, A_1).$$

Chaining this process across a BBN creates the product:

$P(\cap_{k=1}^{n} A_k) = \prod_{k=1}^{n} \; P(A_k | \cap_{j=1}^{k-1} A_j)$, which is a well-known general form of the rule.

For example: $P(A_1 \cap A_2 \cap A_3) = P(A_1 | A_2 \cap A_3) \; P(A_2 | A_3) \; P(A_3)$.

3.3.20 Sample, Sample Space, Random Sample, Simple Random Sample, Random Experiment (Event), and Random Variable

Sample A sample is a collection of units from a population. **Sample Space**. Each trail has as its outcome one of the elements of S, which is the set of all possible outcomes of one single trial, which contains the universe, U, of all possible outcomes. **Random Sample**. A random sample consists of members chosen at random from a given population in such a way that one can compute the chance of obtaining any particular sample. We name the number of units in the sample the sample size, often denoted as n. The number of units in the population often is denoted N. One can draw random samples with or without replacing objects between draws; that is, drawing all n objects in the sample at once (a random sample without replacement), or drawing the objects one at a time, replacing them in the population between draws (a random sample with replacement). In a random sample with replacement, any given member of the population can occur in the sample more than once. In a random sample without replacement, any given member of the population can be in the sample at most once. **Simple Random Sample**. A simple random sample of n units from a population is a random sample drawn by a procedure that is equally likely to give every collection of n units from the population; that is, the probability that the sample will consist of any given subset of n of the N units in the population is $\frac{1}{n} C_n$. Simple random sampling is sampling at random without replacement (without replacing the units between draws). We can construct a simple random sample of size n from a population of $N \geq n$ units by assigning a random number between zero and one to each unit in the population, then taking as the sample those units that this method assigned the n largest random numbers. **Random Experiment (Event)**. An experiment or trial whose outcome is not perfectly predictable, but for which the long-run relative frequency of outcomes of different types in repeated trials is predictable. Note that "random" is different from "haphazard," which does not necessarily imply long-term regularity. **Random Variable**. A random variable is an assignment of numbers to possible outcomes of a random experiment. For example, consider tossing three coins. The number of heads showing when the coins land is a random variable: it assigns the number 0 to the outcome {T, T, T}, the number 1 to the outcome {T, T, H}, the number 2 to the outcome {T, H, H}, and the number 3 to the outcome {H, H, H}.

3.3.21 Real Number

Loosely speaking, real numbers are all numbers that one can represent as fractions (rational numbers), either proper or improper—and all numbers in between the rational numbers. That is, the real numbers comprise the rational numbers and all limits of Cauchy sequences of rational numbers, where the Cauchy sequence is with respect to the absolute value metric. (More formally, the real numbers are the completion of the set of rational numbers in the topology induced by the absolute value function.) The real numbers contain all integers, all fractions, and all irrational (and transcendental) numbers, such as π, e, and 2½. There are uncountably many real numbers between 0 and 1; in contrast, there are only countably many rational numbers between 0 and 1.

3.3.22 Set, Subset, Member of a Set, and Empty Set

Set. A set is a collection of things, without regard to their order. **Subset**. A subset of a given set is a collection of things that belong to the original set. Every element of the subset must belong to the original set, but not every element of the original set need be in a subset (otherwise, a subset would always be identical to its originating set). **Member of a Set**. Something is a member (or element) of a set if it is one of the things in the set. **Empty Set**. The empty set, denoted { } or Ø, is the set that has no members.

3.3.23 Theories of Probability

A theory of probability is a way of assigning meaning to probability statements such as "the chance that a thumb tack lands point-up is 2/3." That is, a theory of probability connects the mathematics of probability, which is the set of consequences of the axioms of probability, with the real world of observation

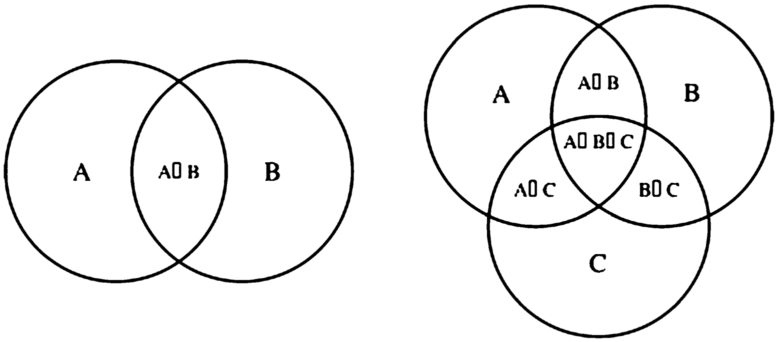

Fig. 3.1 Venn diagram representing (AUB) (*left* Venn diagram) and (AUBUC) (*right* Venn diagram) (Obtained from Venn Diagram (n.d.))

and experiment. There are several common theories of probability. According to the frequency theory of probability, the probability of an event is the limit of the percentage of times that the event occurs in repeated, independent trials under essentially the same circumstances. According to the subjective theory of probability, a probability is a number that measures how strongly we believe an event will occur. The number is on a scale of 0–100%, with 0% indicating that we are completely sure it will not occur, and 100% indicating that we are completely sure that it will occur. According to the theory of equally likely outcomes, if an experiment has *n* possible outcomes, and (for example, by symmetry) there is no reason that any of these outcomes should occur preferentially to any of the others, then the chance of each outcome is 100%/*n*. Each of these theories has its limitations, its proponents, and its detractors.

3.3.24 Unit

A member of a population.

3.3.25 Venn Diagram

A Venn diagram shows the relations among sets or events using diagrams or pictures. We usually draw the universal set or outcome space as a rectangle where we represent sets as probability regions within this rectangle and the overlapping regions are the intersection of the sets. If the regions do not overlap, then we say the sets are disjoint or mutually exclusive, which represent the union of the sets.

The Venn diagram in Fig. 3.1 is a graphical representation of a Venn diagram.

3.4 The Algebra of Sets

The following are a number of applicable general laws about sets that follow from the definitions of set theoretic operations, subsets, etc.[6] Stoll (1979) offers the following two theorems:

3.4.1 *Theorem 1: For Any Subsets, A, B, & C of a Set U the Following Equations Are Identities*

1.	$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$	Associative law
2.	$A \cup B = B \cup A$	$A \cap B = B \cap A$	Communicative law
3.	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	Distributive law
4.	$A \cup \emptyset = A$	$A \cap U = A$	Identity laws
5.	$A \cup \bar{A} = U$	$A \cap \bar{A} = \emptyset$	Complement law

3.4.2 *Theorem 2: For Any Subsets, A and B of a Set U the Following Equations Are Identities*

6.	If, for all A, $A \cup B = A$, then $B = \emptyset$	If, for all A, $A \cap B = A$, then $B = U$	
7.	$A \cup B = U$ and $A \cap B = \emptyset$, then $B = \bar{A}$		Self-dual
8.	$\bar{\bar{A}} = A$		Self-dual
9.	$\bar{\emptyset} = U$	$U = \emptyset$	Identity law
10.	$A \cup A = A$	$A \cap A = A$	Idempotent law
11.	$A \cup U = U$	$A \cap \emptyset = \emptyset$	Identity law
12.	$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	Absorption law
13.	$\overline{A \cup B} = \bar{A} \cap \bar{B}$	$\overline{A \cap B} = \bar{A} \cup \bar{B}$	DeMorgan law

Acknowledgments I would like to thank Dr. Philip B. Stark, Department of Statistics, University of California, Berkeley for allowing me to use his statistical terminology and definitions liberally from his website, SticiGui, found at: http://statistics.berkeley.edu/~stark/SticiGui/index.htm

[6] See Stoll (1979) for the proofs to these identities.

References

Bolstad, W. M. (2007). *Introduction to Bayesian statistics* (2nd ed.). Hoboken, NJ: Wiley.

Hebert, S., Lee, V., Morabito, M., & Polan, J. (2007). *Bayesian network theory*. https://controls.engin.umich.edu/wiki/index.php/Bayesian_network_theory. Accessed 23 July 2012.

Stark, P. B. (2012). *Glossary of statistical terms* http://statistics.berkeley.edu/~stark/SticiGui/Text/gloss.htm. Accessed 26 June 2012.

Stoll, R. R. (1979). *Set theory and logic*. Mineola, NY: Dover Publications.

Venn Diagram (n.d.). http://www.wolframalpha.com/input/?i=venn+diagram. Accessed 26 June 2012.

Chapter 4
Bayesian Belief Networks (BBN) Experimental Protocol

4.1 Introduction

This chapter represents the statistical methodology I followed in formatting the example Chaps. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14 that follow. My intent is to present a starting point reference guide for naïve researchers when exploring and learning BBN.

4.2 BBN Experimental Protocol

This chapter provides a road map to the required statistical and probability theory review to cover the multi-dimensional and hierarchical relationships that exist in BBN.

4.3 Characteristics of a Random Experiment

The follow-on chapters contain ten random example experiments across multiple areas of research interest. Using the statistical definitions defined in Chap. 3, *Statistical Properties of Bayes' Theorem*, each random experiment contains two or more events and elements where I derived the data from random sampling techniques using Monte Carlo simulations. These examples are fictitious in nature but do represent reality at a higher levels of thought and the techniques and procedures I outline here can be easily adapted to similar experiments. To do this, it is critical that researchers understand the premise of conducting the random experiment so they can be grounded in theory. This chapter suggests a methodology as a starting point for the follow-on chapters and in conducting subsequent independent research.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_4,

4.4 Bayes' Research Methodology

There exists a universe of events that have actually occurred in the past that are invisible to us. We have not seen them but we know they exist. Having a scientific mindset, we desire to determine an acceptable truth of the proportions that the elements of these respective events represent. We begin this process by making an initial assumption of these proportions based on our beliefs. We then conduct an experiment by making random draws from a population of interest to determine these proportions. In the language of Bayes' statistics, we refer to this population base as the universe of all possible events in which we seek to make their respective elements visible or known. In the Bayesian universe, we seek to determine conditional relationships across these events. In essence the unobserved event then becomes our "Cause" and the observable events our "Effect" and vice versa.

In general, in doing this we seek to identify a universal set. Within this universe, we also seek mutually exclusive (disjoint) sub-sets. The unobservable sub-set contains the elements we desire to discover. We discover these elements by conditioning the observable on the unobservable event(s) and count these frequencies of dependencies. In essence, we are asking what is the probability of an Event B given the evidence of an Event A, or $\mathrm{P}(\mathrm{B}|\mathrm{A})$ in a BBN. Understanding this concept is critical so that we can properly conduct these experiments.

For example, if we have a vat of white and black marbles and we desire to know the proportional mix of each, then we could conduct a random experiment starting with our initial belief of the mix. Let's say out belief was that 10% of the marbles are black. We would conduct our random draw and select a marble at random while blind folded. This way we do not see the marble we select initially and in theory, one out of every ten marbles is black. We take a series of random draws and beginning with the first marble we select, we show this marble to a spotter who identifies its true color. We are expecting the first marble to be black and the spotter will verify (truthify) its color and then categorize it accordingly and begin the counting process using conditional probability. Let's say we make 100 random draws and 95% of every one out of ten marbles were actually white, $\mathrm{P}(\tilde{W}|\mathrm{W})$, and 5.0% were actually black, $\mathrm{P}(\tilde{B}|\mathrm{W})$, and of the remaining nine out of ten marbles, 80% were actually black, $\mathrm{P}(\tilde{B}|\mathrm{B})$ and 20% were actually white, or $\mathrm{P}(\tilde{W}|\mathrm{B})$. Here, we have set us our random experiment to use for the counting process for Bayes' theorem.[1] Using the statistical properties of Bayes' theorem developed in Chap. 3, I will follow these steps:

[1] This is the random draw method I used in the follow-on example chapters, Chaps. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14, in the book. I explain the symbology embedded in this example in detail below.

4.5 Conducting a Bayesian Experiment

This section outlines the research methodology an analyst can follow when conducting BBN research. I framed it using the following 11-Steps: (1) Step 1: Identify a population of interest, (2) Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population, (3) Step 3: Determine prior (a priori) or unconditional probabilities, (4) Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements, (5) Step 5: Conduct the random experiment, (6) Step 6: Determine Frequency Counts, (7) Step 7: Determine Likelihood/Conditional Probabilities (Relative Frequencies), (8) Step 8: Determine Joint Probabilities, (9) Step 9: Determine Posterior Probabilities, (10) Step 10: Draw a Tree Diagram, and (11) Step 11: Run a Netica Replication. These steps follow iteratively:

Step 1: Identify a population of interest. This is a collection of studied units. Units can be people, places, objects, epochs, drugs, procedures, or many other things. Much of statistics is concerned with estimating numerical properties (parameters) of an entire population from a random sample of units from the population.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) elements, which are the subsets of our population. This represents the disjoint elements of the unobservable event.

Step 3: Determine prior[2] *(a-priori) or unconditional probabilities*. An analyst determines these discrete event probabilities before conducting the experiment to satisfy the independence requirement of Bayes' theorem. This is a common fallacy in the literature and researchers should not use priors they obtain from observable events. The concept of prior information and unconditional[3] probabilities is unique and it represents how confident an analyst is in her or his initial beliefs. These priors are subjective and represent the proportions that she or he believe exist in the population and an analyst weight them before the experiment without looking at the observable data counts and probabilities. I designate them using percentages as: $\mathrm{P(B)} = \mathrm{X}\%$ and $\mathrm{P}(\tilde{\mathrm{B}}) = 1 - \mathrm{P(B)}$, where $\mathrm{P(B)} + \mathrm{P}(\tilde{\mathrm{B}}) = 100\,\%$, where, again, Event B_i represents the unobservable event.[4]

---

[2] These priors can also be vague but we will see that they can be "washed them away" across BNN nodes using the chain rule in Chap. 14, "*Special Forces Assessment and Selection (SFAS) Two Stage Example*." They are really just starting points in our experimental process but must always remain independent. When we start referencing to parameter values and prior distributions, we quickly slip into Bayesian inference techniques such as WinBugs, R, BayesiaLab, and other MCMC software platforms.

[3] See Earman (1992) for a thorough discussion (including many historical references) of the problem of priors and other foundational controversies in Bayesian philosophy of science.

[4] When the observable Event B has more than two elements, $B_1 + B_2 + \cdots + B_n = 100\%$.

Table 4.1 Frequency counts

	Observable event		
Unobservable event	Event A	Event Ã	Total
Event B	A|B[a]	Ã|B	A|B + Ã|B
Event B̃	A|B̃	Ã|B̃	A|B̃+Ã|B̃
Conditional/marginal probabilities			
Total	A|B+A|B̃	Ã|B+Ã|B̃	A|B+A|B̃+Ã|B+Ã|B̃

Note: These values represent the total conditional counts obtained from the sampling process. This is a precursor to computing the likelihood probabilities. An analyst calculates totals by adding across the observable and down the unobservable event. For example, if the count of A|B = 25 and Ã|B = 75, then the total = 100

[a]A|B represents the observable count of event A given an event B

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, an analyst is looking for an observed event consisting of elements she or he can slice through to identify their proportions of outcomes.

Step 5: Conduct the random experiment. In conducting the experiment, the analyst selects at random one of the elements of the unobservable[5] event, Event B_i, based on the a-priori probabilities identified in Step 3. Then, depending on the drawn element, either B or $\tilde{B}$, she or he will take a random sample from the observable event, Event A_i, to select either element A or $\tilde{A}$ and assign this conditionally to obtain their initial count. This will be done iteratively until they have obtained their desired sample size.

Step 6: Determine frequency counts. To record frequency counts, an analyst will report these in a Frequency Count matrix as illustrated in Table 4.1. After each repetition of the experiment, the analyst will sum across the observable and down the unobservable event to determine totals.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood probabilities as conditional based on the subjective probabilities of the unobservable event. To determine these percentages, the analyst calculates probabilities across the sliced elements of the observable event and then reports these results in Table 4.2, Relative Frequency/ Likelihood/Conditional Probabilities:

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 4.2 across the observable and unobservable events. To compute conditional (marginal) probabilities, the analyst then sums the joint probabilities down the elements of the unobservable event and then across the elements of the observable event, which totals 100.0%. The analyst then reports these in Table 4.3 as joint and marginal probabilities:

[5] Since this is an unobservable event, the analyst cannot identify the elements of this set so the conditional identification of the observable element with it allows for conduct of the counting.

Table 4.2 Relative frequency/likelihood/conditional probabilities

Unobservable event	Observable event P(A)	P(Ã)	Total (%)
P(B)	P(A|B)[a]	$P(\tilde{A}\|B)$	100.0
$P(\tilde{B})$	$P(A\|\tilde{B})$	$P(\tilde{A}\|\tilde{B})$	100.0
Conditional/marginal probabilities			
Total	$P(A\|B)+P(A\|\tilde{B})$[b]	$P(A\|\tilde{B})+P(\tilde{A}\|\tilde{B})$	100.0

Note: These values represent relative frequencies/likelihood probabilities for the conditional sampling that the analyst calculated using count data reported in Table 4.1. The analyst computes the conditional/marginal probabilities by dividing the total frequency counts down the observable event and across unobservable event using the frequency counts from Table 4.1

[a]$P(A|B) = A|B/(A|B + \tilde{A}|B) \times 100$

[b]$P(A|B) + P(A|\tilde{B}) = (A|B + A|\tilde{B})/\ (A|B + A|\tilde{B} + \tilde{A}|B + \tilde{A}|\tilde{B}) \times 100.$

Table 4.3 Joint and marginal probabilities

Unobservable event	Observable event P(A)	P(Ã)	Marginal probabilities
P(B)	$P(A\cap B)$[a] $= P(A\|B)P(B)$	$P(\tilde{A}\cap B) = (P(\tilde{A}\|B)P(B)$	$P(A \cap B) + P(\tilde{A}\cap B)$
$P(\tilde{B})$	$P(A\cap\tilde{B}) = P(A\|\tilde{B})P(\tilde{B})$	$P(\tilde{A}\cap\tilde{B}) = P(\tilde{A}\|\tilde{B})P(\tilde{B})$	$P(A\cap\tilde{B}) = P(\tilde{A}\cap\tilde{B})$
Conditional (marginal) probabilities			
Total	$P(A\cap B)+P(A\cap\tilde{B})$	$P(\tilde{A}\cap B)+P(\tilde{A}\cap\tilde{B})$	100%

Notes: These values represent the joint probabilities for each observable and unobservable event elements that the analyst calculated using prior probabilities and probabilities reported in Table 4.2, Relative Frequency/Likelihood/Conditional Probabilities

[a]$P(A\cap B)$ = Prior probability × P(A|B). The observable and unobservable events are dependent if $P(A\cap B) \neq [P(A\cap B) + P(A\cap\tilde{B}) \times P(A\cap B) + P(\tilde{A}\cap B)]$

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 4.3 by their respective conditional/marginal probabilities, which totals 100%. For example, the analyst computes the posterior probabilities for each element in the observable event by dividing them individually by their respective conditional/marginal probabilities and then reports these in Table 4.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 4.1.

Table 4.4 Posterior probabilities

	Observable event	
Unobservable event	P(A)	$P(\tilde{A})$
P(B)	$P(A \cap B)/P(A)$[a]	$P(A \cap \tilde{B})/P(\tilde{A})$
$P(\tilde{B})$	$P(\tilde{A} \cap B)/P(A)$	$P(\tilde{A} \cap \tilde{B})/P(\tilde{A})$
Total	$P(A \cap B)/P(A) + P(\tilde{A} \cap B)/P(A) = 100\%$	$P(A \cap \tilde{B})/P(\tilde{A}) + P(\tilde{A} \cap \tilde{B})/P(\tilde{A}) = 100\%$

Note: This represents the posterior probabilities of the elements of the observable event. The analyst calculated them using the joint and conditional/marginal probabilities reported in Table 4.3

[a]$P(A \cap B)/P(A) = P(A \cap B)/[P(A \cap B) + P(A \cap \tilde{B})]$

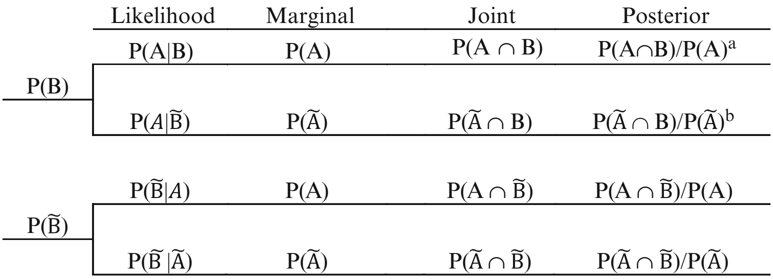

Fig. 4.1 Tree diagram. From Tables 4.1 through 4.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model. [a]$P(A) = P(B \cap A)\ P(A) + P(B \cap \tilde{A})\ P(\tilde{A})$ (marginal probabilities). [b]$P(\tilde{A}) = P(B \cap A)\ P(A) + P(B \cap \tilde{A})\ P(\tilde{A})$(marginal probabilities)

Step 11: Run a netica replication.[6] The analyst reports the results of the Netica (The Norsys Software Corp[7]) replication of the prior, conditional, and marginal probabilities of the BBN.

Reference

Earman, J. (1992). *Bayes or bust: A critical examination of Bayesian confirmation theory*. Cambridge: MIT Press.

[6] By default, Netica only adds one decimal point for two-digit numbers and two decimal points for one-digit numbers. To compensate for this, I calculated all numbers with one decimal point. Note, there may be some rounding errors due to this.

[7] See http://www.norsys.com/

Chapter 5
Manufacturing Example*

5.1 Scenario

In this scenario, XYZ Electronics, Inc. obtains transistors from multiple companies and stores them in an open container. The assembly department is experiencing an above-average number of defective transistors from the deliveries of three suppliers, Companies A, B, and C. Their concern is that these defective parts will start slowing down assembly time in the plant and exponentially increase the cost of goods sold. Their research question is to determine the proportions of defective and non-defective transistors given the contributions of respective suppliers, Company A, B, and C. Obtaining quality transistors with the minimal amount of costs would be a benefit to the XYZ Electronics, Inc. A company analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

5.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of transistors produced and delivered by each of these three companies.

Step 2: slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The two disjoint elements are "Not Defective" and "Defective" transistors from the element, "Transistor Quality."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, ABC Electronics, Inc. has received 28.6% defective transistors from their suppliers.

---

*This context of this example is from Weiers (Weiers et al. 2005).

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_5,

Table 5.1 Frequency counts

Transistor quality	Supplier Company A	Company B	Company C	Total
Not defective	120	83	72	275[a]
Defective	17	12	31	60
Total	137[b]	95	103	335

Note: These values represent transistor quality frequency counts for each of the Supplier elements
[a]275 = 120 + 83 + 72
[b]137 = 120 + 17

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that has the capability to identify which company is supplying defective transistors. The disjoint event is "Supplier." The analyst will slice through Supplier by identifying the effects following the distribution of capacitors from each of the three companies: (1) Company A, (2) Company B, and (3) Company C, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of Good and Defective transistors inside a storage container. The sampling process starts with a single random draw and selection of an element from Transistor Quality and then from Supplier and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 5.1, Frequency Counts, for 335 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 5.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 5.2 across Supplier and Transistor Quality. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Transistor Quality and then across the elements of Supplier, which totals 100.0%. The analyst then reports these in Table 5.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 5.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Supplier by dividing them individually by their respective marginal probabilities and then reports these in Table 5.4 as posterior probabilities.

Table 5.2 Relative frequency/likelihood/conditional probabilities

Transistor quality	Supplier			
	Company A (%)	Company B (%)	Company C (%)	Total (%)
Not defective	43.6[a]	30.2	26.2	100.0
Defective	28.3	20.0	51.7	100.0
Conditional probabilities				
Total	40.9[b]	28.4	30.7	100.0

Note: These values represent transistor quality relative frequencies/likelihood/conditional probabilities for each Supplier that the analyst calculated using count data reported in Table 5.1

[a]43.6% = 120/275 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Transistor Quality and across Supplier using the frequency counts from Table 5.1

[b]40.9% = 137/335 × 100

Table 5.3 Joint and marginal probabilities

Transistor quality	Supplier			Marginal probabilities (%)
	Company A (%)	Company B (%)	Company C (%)	
Not defective	31.2[a]	21.6	18.7	71.4[c]
Defective	8.1	5.7	14.7	28.6
Marginal probabilities				
Total	39.3[b]	27.3	33.5	100.0

Notes: These values represent the joint probabilities for each Transistor Quality and Supplier elements that the analyst calculated using prior probabilities and probabilities reported in Table 5.2, Relative Frequency/Likelihood Probabilities. Events Supplier and Transistor Quality are dependent as evaluated by P(Company A ∩ Not Defective) ≠ P(Company A) × P(Not Defective), 31.2% ≠ 39.3% × 71.4% = 28.1%

[a]31.2% = 71.4% × 43.6%

[b]39.3% = 31.2% + 8.1%

[c]71.4% = 31.2% + 21.6% + 18.7%

Table 5.4 Posterior probabilities

Transistor quality	Supplier		
	Company A (%)	Company B (%)	Company C (%)
Not defective	79.4[a]	79.0	55.9
Defective	20.6	21.0	44.1
Total	100.0[b]	100.0	100.0

Note: This represents the posterior probabilities of the elements of Supplier. The analyst calculated them using the joint and marginal probabilities reported in Table 5.3

[a]79.4% = 31.2%/39.3%

[b]100.0% = 79.4% + 20.6%

Manufacturing Example						
Transistor Quality		Company	Likelihood	Joint	Marginal	Posterior
Not Defective	71.4%	Company A	43.6%	31.2%	39.3%	79.4%
		Company B	30.2%	21.6%	27.3%	79.0%
		Company C	26.2%	18.7%	33.5%	55.9%
Defective	28.6%	Company A	28.3%	8.1%		20.6%
		Company B	20.0%	5.7%		21.0%
		Company C	51.7%	14.8%		44.1%

Fig. 5.1 Tree diagram for the manufacturing example. From Tables 5.1, 5.2, 5.3, and 5.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this one-stage BBN model

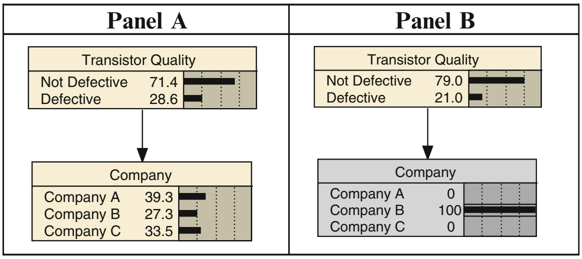

Fig. 5.2 This represents the Netica replication of this manufacturing example. *Panel A* represents the subjective or prior probabilities (transistor quality node) and the conditional or marginal joint probabilities (company node) (Table 5.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Company B (Table 5.4). The analyst uses the raw data inputs into the Netica un-normalized input table

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 5.1.

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 5.2.

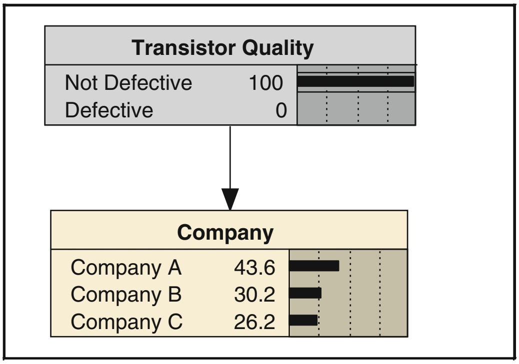

Fig. 5.3 Represents the effects on the conditional probabilities using inverse probability when the analyst inverts the cause and effect relationship. For example in referring to Fig. 5.3, not only can the analyst determine the probability of a unloaded dice effecting a gambler's win or losing, P(Not Defective|Company B) = 79.0 % (Fig. 5.2), but they can also determine the percentage effect of not detective transistors from Company B (Company A and Company C), P(Company B|Defective) = 30.2 %, which are two distinct probabilities. Invoking Company or Company B will similarly adjust these conditional probabilities

5.3 Conclusions

After conducting this experiment, the analyst has a BBN loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

5.3.1 Posterior Probabilities

The analyst can evaluate conditional probability changes going from the effect event, Transistor Quality, to the cause event, Company. After conducting this experiment and priming the BBN with all available information, following the outcome of the next Company B produced transistor, she or he could conclude that there is a 79.0% chance that it is not defective and a 21.0% chance that it is defective. Other events that could contribute to the updating of the posterior probabilities of this BBN include manufacturing quality and standard, number of parts produced, and increases in materials or labor costs.

5.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Transistor Quality, to the cause event, Company. For example in referring to Fig. 5.3, not only can she or he determine the probability of a Unloaded Dice effecting a gambler's win or losing, P(Not Defective|Company B) = 79.0% (Fig. 5.2), but they can also determine the percentage effect of Not defective transistors from Company B (Company A and Company C), P(Company B|Defective) = 30.2%, which are two distinct probabilities. Invoking Defective will similarly adjust these conditional probabilities.

References

Weiers, R. M., Gray, B. J., & Peters, L. H. (2005). *Introduction to business statistics* (5th ed.). Australia/Mason, OH: South-Western Cengage Learning.

Chapter 6
Political Science Example

6.1 Scenario

In this scenario, an incumbent State Republican Senator obtains constituent votes from two majority Republican counties and hopes to maintain voter confidence. The Senator is experiencing an above-average number of party defectors from these two counties, County A and B. The Senator's concern is that these defectors will sway other voters and jeopardize the upcoming elections and exponentially increase the cost of winning these voters back. The Senator's research question is to determine the proportions of remaining current political affiliation. Obtaining loyal voters with the minimal amount of costs would be a benefit to the winning strategy of the campaign. A political analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

6.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of registered voters in each of the concerned counties.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The two disjoint elements are "County A" and "County B" from the element, "County."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the 28.6% of the constituents in County A and 71.4% of County B have supported the incumbent.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that can identify county constituency. The disjoint event is, "Constituency."

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_6,

Table 6.1 Frequency counts

County	Political affiliation		
	Democrat	Republican	Total
County A	243	145	388[a]
County B	211	151	362
Total	454[b]	296	750

Note: These values represent county counts for each of the Political Affiliation elements
[a]388 = 243 + 145
[b]454 = 243 + 211

Table 6.2 Relative frequency/likelihood/conditional probabilities

County	Political affiliation		
	Democrat (%)	Republican (%)	Total (%)
County A	62.6[a]	37.4	100.0
County B	58.3	41.7	100.0
Conditional probabilities			
Total	60.5[b]	39.5	100.0

Note: These values represent county relative frequencies/likelihood/conditional probabilities for each political affiliation that the analyst calculated using count data reported in Table 6.1
[a]62.6 % = 120/275 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down County and across Political Affiliation using the frequency counts from Table 6.1
[b]60.5 % = 454/750 × 100

The analyst will slice through this event by identifying the effects following the results of a survey that will identify these county citizens as either a Democrat or Republican, which will become the sub-events for this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of constituents from Counties A and B in a database. The sampling process starts with a single random draw and selection of an element from County and then from Political Affiliation and ends with the assignment of draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 6.1, Frequency Counts, for 388 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 6.2, Relative Frequency/Likelihood/Conditional Probabilities.

Table 6.3 Joint and marginal probabilities

County	Political affiliation		
	Democrat (%)	Republican (%)	Marginal probabilities (%)
County A	17.9[a]	10.7	28.6[c]
County B	41.6	29.8	71.4
Marginal probabilities			
Total	59.5[b]	40.5	100.0

Notes: These values represent the joint probabilities for each Country and Political Affiliation elements that the analyst calculated using prior probabilities and probabilities reported in Table 6.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Political Affiliation and County are dependent as evaluated by P(Democrat ∩ County A) ≠ P(Democrat) × P(County A), 17.9% ≠ 59.5% × 28.6% = 17.0%
[a]17.9% = 28.6% × 62.6%
[b]59.5% = 17.9% + 41.6%
[c]28.6% = 17.9% + 10.7%

Table 6.4 Posterior probabilities

County	Political affiliation	
	Democrat (%)	Republican (%)
County A	30.1[a]	26.4
County B	69.9	73.6
Total	100.0[b]	100.0

Note: This represents the posterior probabilities of the elements of Political Affiliation. The analyst calculated them using the joint and marginal probabilities reported in Table 6.3
[a]30.1% = 17.9%/59.5%
[b]100.0% = 30.1% + 69.9%

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 6.2 across Political Affiliation and County. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of County and Political Affiliation, which totals 100.0%. The analyst then reports these in Table 6.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 6.4 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Political Affiliation by dividing them individually by their respective marginal probabilities and then reports these in Table 6.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 6.1.

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 6.2.

Political Science Example						
County		Political Affiliation	Likelihood	Joint	Marginal	Posterior
County A	28.6%	Democrat	62.6%	17.9%	59.5%	30.1%
		Republican	37.4%	10.7%		26.4%
County B	71.4%	Democrat	58.3%	41.6%		69.9%
		Republican	41.7%	29.8%	40.5%	73.6%

Fig. 6.1 Tree diagram for the Political Science example. From Tables 6.1, 6.2, 6.3 and 6.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

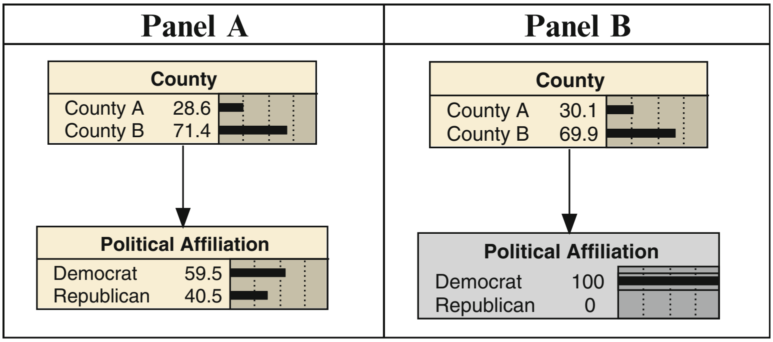

Fig. 6.2 This represents the Netica replication of this Political Science example. *Panel A* represents the subjective or prior probabilities (County Node) and the conditional or marginal joint probabilities (Political Affiliation Node) (Table 6.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Democrat (Table 6.4). The analyst uses the raw data inputs into the Netica un-normalized input table

6.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

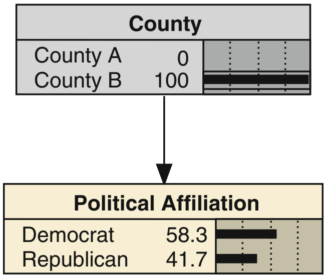

Fig. 6.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 6.3, not only can an analyst determine the probability of democrats living in County B, P(County B|Democrat) = 69.9% (Fig. 6.2), but she or he can also determine the percentage effect of democrats (and republicans) living in County B, P(Democrat|County B) = 58.3%, which are two distinct probabilities. Invoking County A will similarly adjust these conditional probabilities

6.3.1 Posterior Probabilities

Here, an analyst can evaluate conditional probability changes going from the effect event, County, to the cause event, Political Affiliation. After conducting this experiment and priming the BBN with all available information, following the outcome of the next citizen polling and identifying a Democrat, an analyst could conclude that there is a 30.1% chance that she or he was from County A and a 69.9% chance that she or he was from County B. Other events that could contribute to updating the posterior probabilities of this BBN could include length of time of their current political affiliation and conservative or liberal political views.

6.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, County, to the cause event, Political Affiliation. For example in referring to Fig. 6.3, not only can an analyst determine the probability of democrats living in County B, P(County B|Democrat) = 69.9% (Fig. 6.2), but she or he can also determine the percentage effect of democrats(and republicans) living in County B, P(Democrat|County B) = 58.3%, which are two distinct probabilities. Invoking County A will similarly adjust these conditional probabilities.

Chapter 7
Gambling Example

7.1 Scenario

In this scenario, a casino obtains die from a custom manufacturer who ensures fairness and issues them to employees working at tables that use these die. The casino experiencing an above-average number of wins from its gamblers who are either using loaded or unloaded die. Their concern is that loaded die will start slowing eroding away the advantage of the house and increase the loss rate of the casino. Their research question is to determine the proportions of die fairness given the contributions of respective die. Obtaining fair die with the minimal amount of costs would be a benefit to the casino. A casino analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

7.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of all die that have been thrown in the casino by all gamblers.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The two disjoint elements are "Winner" and "Loser" gamblers from the element, "Die Randomness."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the casino has experienced 50.0% of gamblers who normally win.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that can determine which gamblers are using loaded die. The disjoint event is, "Fair Die." The analyst will slice through Fair Die by identifying the

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_7, © Springer Science+Business Media New York 2013

Table 7.1 Frequency counts

Die randomness	Fair die Loaded	Unloaded	Total
Winner	4	7	11[a]
Loser	58	49	107
Total	62[b]	56	118

Note: These values represent die randomness frequency counts for each of the Fair Die elements
[a]11 = 4 + 7
[b]62 = 4 + 58

Table 7.2 Relative frequency/likelihood/conditional probabilities

Die randomness	Fair die Loaded (%)	Unloaded (%)	Total (%)
Winner	36.4[a]	63.6	100.0
Loser	54.2	45.8	100.0
Conditional probabilities			
Total	52.5[b]	47.5	100.0

Note: These values represent die randomness relative frequencies/likelihood/conditional probabilities for each of the fair die that the analyst calculated using count data reported in Table 7.1
[a]36.4% = 4/11 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Transistor Quality and across Supplier using the frequency counts from Table 7.1
[b]52.5% = 62/118 × 100

effects following the play of Loaded and Unloaded Die from each of the gamblers, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of winners and loser inside a database. The sampling process starts with a single random draw and selection of a sub-event from Die Randomness and then from Fair Die and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 7.1, Frequency Counts, for 11 iterations.

*Step 7: Determine likelihood/conditional probabilities (relative frequencies).*The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 7.2, Relative Frequency/Likelihood/Conditional Probabilities.

Table 7.3 Joint and marginal probabilities

	Fair die		
Die randomness	Loaded (%)	Unloaded (%)	Marginal probabilities (%)
Winner	18.2[a]	31.8	50.0[c]
Loser	27.1	22.9	50.0
Marginal probabilities			
Total	45.3[b]	54.7	100.0

Notes: These values represent the joint probabilities for each Transistor Quality and Supplier elements that the analyst calculated using prior probabilities and probabilities reported in Table 7.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Die Status and Die Randomness are dependent as evaluated by P(Loaded $\cap$ Winner) $\neq$ P(Loaded) $\times$ P(Winner), 18.2%$\neq$ 45.3% $\times$ 50.0% = 22.6%

[a]18.2% = 50.0% $\times$ 36.4%

[b]45.3% = 18.2% + 27.1%

[c]50.0% = 18.2% + 31.8%

Table 7.4 Posterior probabilities

	Fair die	
Die randomness	Loaded (%)	Unloaded (%)
Winner	40.2[a]	58.2
Loser	59.8	41.8
Total	100.0[b]	100.0

Note: This represents the posterior probabilities of the elements of Fair Die. The analyst calculated them using the joint and marginal probabilities reported in Table 7.3

[a]40.2% = 18.2%/45.3%

[b]100.0% = 40.2% + 59.8%

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 7.2 across Die Randomness and Fair Die. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Die Randomness and Fair Die, which totals 100.0%. The analyst then reports these in Table 7.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 7.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Supplier by dividing them individually by their respective marginal probabilities and then reports these in Table 7.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 7.1.

Gambling Example						
Dice Randomness		Fair Die	Likelihood	Joint	Marginal	Posterior
Winner	50.0%	Loaded	36.4%	18.2%	45.3%	40.2%
		Unloaded	63.6%	31.8%	54.7%	58.2%
Loser	50.0%	Loaded	54.2%	27.1%		59.8%
		Unloaded	45.8%	22.9%		41.8%

Fig. 7.1 Tree diagram for the Gambling Example. From Tables 7.1, 7.2, 7.3, and 7.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

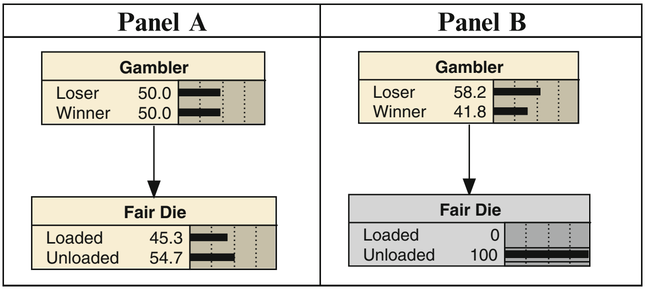

Fig. 7.2 This represents the Netica replication of this Gambling example. *Panel A* represents the subjective or prior probabilities (Gambler Node) and the conditional or marginal joint probabilities (Fair Die Node) (Table 7.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Unloaded (Table 7.4). The analyst uses the raw data inputs into the Netica un-normalized input table

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 7.2.

7.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

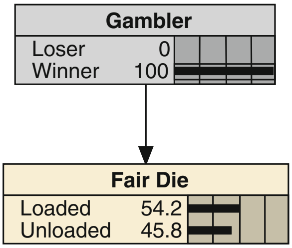

Fig. 7.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 7.3, not only can an analyst determine the probability of Unloaded Die effecting a Winner Gambler, P(Winner| Unloaded) = 41.8% (Fig. 7.2), but she or he can also determine the percentage effect on Unloaded (Loaded) given a Winner Gambler, P(Unloaded|Winner) = 45.8%, which are two distinct probabilities. Invoking Loaded will similarly adjust these conditional probabilities

7.3.1 Posterior Probabilities

Here, an analyst can evaluate conditional probability changes going from the effect event, Gambler, to the cause event, Fair Die. After conducting this experiment and priming the BBN with all available information, following the next gambler with unloaded Die, an analyst could conclude that there is a 58.2% chance that she or he will lose and a 41.8% chance that she or he will win. Other events that could contribute to the updating of the posterior probabilities of this BBN include bankruptcy filing of the gambler, profession, and high-risk personality traits.

7.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Gambler, to the cause event, Fair Die. For example in referring to Fig. 7.3, not only can an analyst determine the probability of Unloaded Die effecting a Winner Gambler, P(Winner|Unloaded) = 41.8% (Fig. 7.2), but she or he can also determine the percentage effect on Unloaded (Loaded) given a Winner Gambler, P(Unloaded|Winner) = 45.8%, which are two distinct probabilities. Invoking Loser will similarly adjust these conditional probabilities.

Chapter 8
Publicly Traded Company Default Example

8.1 Scenario

In this scenario, there is a concern from the Office of the Controller of the Currency (OCC) maintains facility ratings of multinational corporations operating in the U.S. due to potential devaluation issued of the U.S. Dollar (USD) and other global economic issues. The OCC is noticing an above number of companies that they improperly rate that continue as a going concern, are involved. Their concern is the population of U.S. and Non U.S. public multinational companies (MNC) listed on the New York Stock Exchange (NYSE) with high market capitalizations that have great influences on the USD. Their research question is to determine the proportions of Altman Z-Scores ratings they assign to these MNC given they remain as going concerns, transition in merger or acquisition (M&A) activities, are dissolved, or go bankrupt. A U.S. Government OCC analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

8.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of current and historical publicly traded companies listed on the NYSE.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The two disjoint elements are "Z-Scores $>$ 3" and "Z-Scores $< =$ 3" from the element, "Z-Score."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the NYSE has experienced 5.0% of their listed publicly traded companies with Z-Scores $< =$ 3.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_8,

Table 8.1 Frequency counts

	Health status				
Z-Scores	Going concern	M&A	Dissolved	Bankrupt	Total
Z-Scores < = 3	256	583	1,978	239	3,056[a]
Z-Scores >3	6,439	459	321	152	7,371
Total	6,695[b]	1,042	2,299	391	10,427

Note: These values represent z-scores frequency counts for each of the Health Status elements
[a]3,056 = 256 + 583 + 1,978 + 239
[b]6,695 = 256 + 6.439

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that has the capability to identify which company will transition from a going concern. The disjoint event is "Legal Status." The analyst will slice through "Legal Status" by identifying companies that have: (1) continue as a going concern, (2) filed for bankruptcy, (3) entered into an M&A activity, and (4) have simply dissolved, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of companies with Z-Scores > 3 or Z-Scores < = 3 from a database of companies that have been listed on the NYSE. The sampling process starts with a single random draw and selection of a sub-event from Z-Scores and then from Health Status and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 8.1, Frequency Counts, for 10,427 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 8.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 8.2 across Z-Scores and Health Status. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Z-Scores and Health Status, which totals 100.0%. The analyst then reports these in Table 8.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 8.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Health Status by dividing them individually by

Table 8.2 Relative frequency/likelihood/conditional probabilities

	Health status				
Z-Scores	Going concern (%)	M&A (%)	Dissolved (%)	Bankrupt (%)	Total (%)
Z-Scores < = 3	8.4[a]	19.1	64.7	7.8	100.0
Z-Scores >3	87.4	6.2	4.4	2.1	100.0
Conditional probabilities					
Total	64.2[b]	10.0	22.0	3.7	100.0

Note: These values represent Z-Scores relative frequencies/likelihood/conditional probabilities for each Health Status category that the analyst calculated using count data reported in Table 8.1

[a]8.4% = 256/3,056 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Z-Scores and across Health Status using the frequency counts from Table 8.1

[b]64.2% = 6,695/10,427 × 100

Table 8.3 Joint and marginal probabilities

	Health status				
Z-Scores	Going concern (%)	M&A (%)	Dissolved (%)	Bankrupt (%)	Marginal probabilities (%)
Z-Scores < = 3	0.4[a]	1.0	3.2	0.4	5.0[c]
Z-Scores >3	83.0	5.9	4.1	2.0	95.0
Marginal probabilities					
Total	83.4[b]	6.9	7.4	2.4	100.0

Notes: These values represent the joint probabilities for each Z-Scores and Health Status elements that the analyst calculated using prior probabilities and probabilities reported in Table 8.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Health Status and Legal Status are dependent as evaluated by P(Going Concern ∩ Z-Scores > 3) ≠ P(Going Concern) × P(Z-Scores > 3), 0.4% ≠ 83.4% × 5.0% = 4.2%

[a]0.4% = 5.0% × 8.4%

[b]83.4% = 0.4% + 83.0%

[c]5.0% = 0.4% + 1.0% + 3.2% + 0.4%

Table 8.4 Posterior probabilities

	Health status			
Z-Scores	Going concern (%)	M&A (%)	Dissolved (%)	Bankrupt (%)
Z-Scores < = 3	0.5[a]	13.9	43.9	16.6
Z-Scores >3	99.5	86.1	56.1	83.4
Total	100.0[b]	100.0	100.0	100.0

Note: This represents the posterior probabilities of the elements of Health Status. The analyst calculated them using the joint and marginal probabilities reported in Table 8.3

[a]0.5% = 0.4%/83.4%

[b]100.0% = 0.5% + 99.5%

their respective marginal probabilities and then reports these in Table 8.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 8.1.

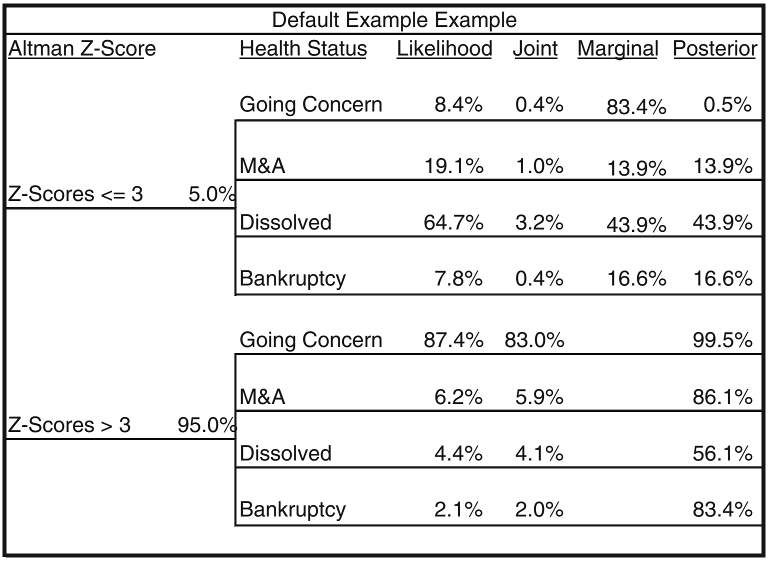

Fig. 8.1 Tree diagram for the Default Example. From Tables 8.1, 8.2, 8.3 and 8.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

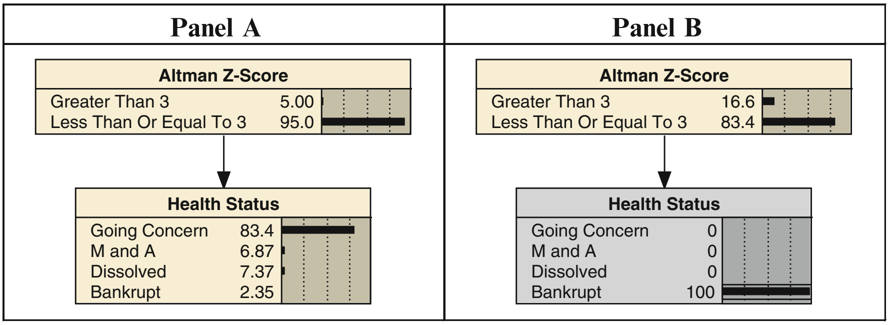

Fig. 8.2 This represents the Netica replication of this default example. *Panel A* represents the subjective or prior probabilities (Altman Z-Score Node) and the conditional or marginal joint probabilities (Health Status Node) (Table 8.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Bankrupt (Table 8.4). The analyst uses the raw data inputs into the Netica un-normalized input table

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 8.2.

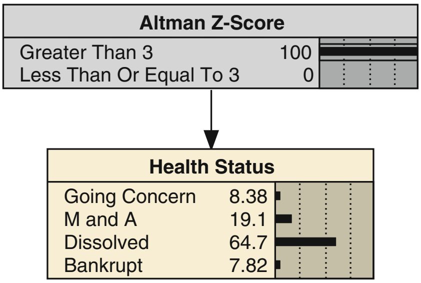

Fig. 8.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 8.3, not only can an analyst determine the probability of a bankruptcy effecting an Altman Z-Score > 3, P(Z-Score > 3|Bankruptcy) = 16.6% (Fig. 8.2), but she or he can also determine the percentage effect on Altman Z-Score > 3 (Z-Score < = 3) given a Bankruptcy, P(Bankruptcy|Z-Score > 3) = 7.82%, which are two distinct probabilities. Invoking Altman Z-Score < = 3 will similarly adjust these conditional probabilities

8.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

8.3.1 Posterior Probabilities

Here, an analyst can evaluate conditional probability changes going from the effect event, Country, to the cause event, Fatality Status. After conducting this experiment and priming the BBN with all available information, following the next company default or bankruptcy, an analyst could conclude there is a 16.6% chance that this company has an Altman Z-Score < 3 and a 83.4% chance an Altman Z-Score > =3. Other events that could contribute to the updating of the posterior probabilities of this BBN include unsuccessful default on debt issues, corporate executive turnover, effects of a recession on the country currency valuation.

8.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Altman Z-Score, to the cause event, Health Status. For example in referring to Fig. 8.3, not only can an analyst determine the probability of a bankruptcy effecting an Altman Z-Score > 3, P(Z-Score > 3|Bankruptcy) $= 16.6\%$ (Fig. 8.2), but she or he can also determine the percentage effect on Altman Z-Score > 3 (Z-Score $< = 3$) given a Bankruptcy, P(Bankruptcy|Z-Score > 3) $= 7.82\%$, which are two distinct probabilities. Invoking Altman Z-Score $< = 3$ will similarly adjust these conditional probabilities.

Chapter 9
Insurance Risk Levels Example

9.1 Scenario

In this scenario, an insurance company obtains insurer risk categories for multiple fatality status and provides insurers with respective rate quotes. The agency is experiencing an above-average number of discounted quotes that do not hedge the risk of respective insurers. Their concern is that these discounted quotes will start putting pressure on general policy holders to make up for forecasted pay outs to fatality insurers and increasing the cost of policies sold exponentially. Their research question is to determine the proportions of age group insurance premiums given the results of future fatality statuses. Obtaining optimal insurance pricing with the minimal amount of costs would be a benefit to the insurance company and its agents. An insurance actuary will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

9.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of drivers that ABC Insurance Company insures.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The three disjoint elements are: "16–19," "20–25," and "$>$ 25" from the element, "Risk Category."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the insurance company has reported fatality rates of 45.0% for Age Group 16–19, 35.0% for Age Group 20–25, and 20.0% for Age Group $>$ 25 of their insurers.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_9,

Table 9.1 Frequency counts

Risk category	Fatality status: Fatality	Fatality status: No fatality	Total
Age 16–19	269	838	1,107[a]
Age 20–24	273	649	922
Age > 25	59	561	620
Total	601[b]	2,048	2,649

Note: These values represent risk category frequency counts for each of the Fatality Status elements
[a]1,107 = 269 + 838
[b]601 = 269 + 273 + 59

an outcome that can identify the future fatality status of an insurer. The disjoint event is, "Fatality Status." The analyst will slice through "Fatality Status" by identifying the effects following speeding incidents of insurer that result in a "Fatality" or "Non-Fatality," which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of their insured drivers in brackets of: (1) Age 16–19, (2) Age 20–24, and (3) Age > 25 from a database. The sampling process starts with a single random draw and selection of an element from Risk Category and then from Fatality Status and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 9.1, Frequency Counts, for 1,107 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 9.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 9.2 across Fatality Status and Risk Category. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Risk Category and Fatality Status, which totals 100.0%. The analyst then reports these in Table 9.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 9.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Fatality Status by dividing them individually by

Table 9.2 Relative frequency/likelihood/conditional probabilities

	Fatality status		Total
Risk category	Fatality (%)	No fatality (%)	Total (%)
Age 16–19	24.3[a]	75.7	100.0
Age 20–24	29.6	70.4	100.0
Age > 25	9.5	90.5	100.0
Total	24.3	75.7	100.0
Conditional probabilities			
Total	22.7[b]	77.3	100.0

Note: These values Risk Category relative frequencies/likelihood/conditional probabilities for each Fatality Status category that the analyst calculated using count data reported in Table 9.1

[a]24.3% = 269/1,107 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Risk Category and across Fatality Status using the frequency counts from Table 9.1

[b]22.7% = 601/2,649 × 100

Table 9.3 Joint and marginal probabilities

	Fatality status		
Risk category	Fatality (%)	No fatality (%)	Marginal probabilities (%)
Age 16–19	10.9[a]	34.1	45.0[c]
Age 20–24	10.4	24.6	35.0
Age > 25	1.9	18.1	80.0
Marginal probabilities			
Total	23.2[b]	76.8	100.0

Notes: These values represent the joint probabilities for each Risk Category and Fatality Status elements that the analyst calculated using prior probabilities and probabilities reported in Table 9.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Fatality Status and Risk Category are dependent as evaluated by P(Fatality ∩ Age 16–19) ≠ P(Fatality) × P(Age 16), 10.9%≠ 23.2% × 45.0% = 10.4%

[a]10.9% = 45.0% × 24.3%

[b]23.2% = 10.9% + 10.4% + 1.9%

[c]45.0% = 10.9% + 34.1%

Table 9.4 Posterior probabilities

	Fatality status	
Risk category	Fatality (%)	No fatality (%)
Age 16–19	47.1[a]	44.4
Age 20–24	44.7	32.1
Age > 25	8.2	23.6
Total	100.0[b]	100.0

Note: This represents the posterior probabilities of the elements of Fatality Status. The analyst calculated them using the joint and marginal probabilities reported in Table 9.3

[a]47.1% = 10.9%/23.2%

[b]100.0% = 47.1% + 44.7% + 8.2%

Insurance Risk Levels Example

Risk Category		Fatality Status	Likelihood	Joint	Marginal	Posterior
Age 16-19	45.0%	Fatality	24.3%	10.9%	23.2%	47.1%
		No Fatality	75.7%	34.1%	76.8%	44.4%
Age 20-24	35.0%	Fatality Status	29.6%	10.4%		44.7%
		No Fatality	70.4%	24.6%		32.1%
Age > 25	20.0%	Fatality	9.5%	1.9%		8.2%
		No Fatality	90.5%	18.1%		23.6%

Fig. 9.1 Tree diagram for the Insurance Risk Levels example. From Tables 9.1, 9.2, 9.3 and 9.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

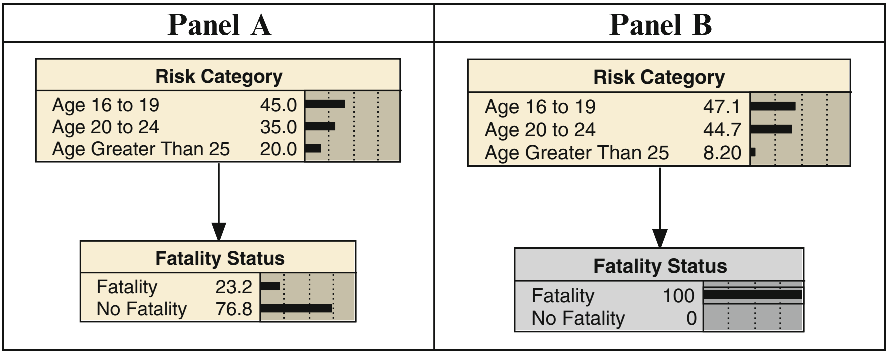

Fig. 9.2 This represents the Netica replication of this Insurance example. *Panel A* represents the subjective or prior probabilities (Risk Category Node) and the conditional or marginal joint probabilities (Fatality Status Node) (Table 9.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Fatality (Table 9.4). The analyst uses the raw data inputs into the Netica un-normalized input table

their respective marginal probabilities and then reports these in Table 9.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 9.1.

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 9.2.

9.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

9.3.1 Posterior Probabilities

Here, an analyst can evaluate conditional probability changes going from the effect event, Risk Category, to the cause event, Fatality Status. After conducting this experiment and priming the BBN with all available information, following the next fatal automobile accident, an analyst could conclude there is a 47.1% chance that this occurred in the Age 16–19 age group, a 44.7% chance in the Age 20–24 age group, and a 8.2% chance in the Age > 25 age group. Other events that could contribute to the updating of the posterior probabilities of this BBN include previous arrest due to driving while intoxication, previous driving records, and other risky lifestyle indicators.

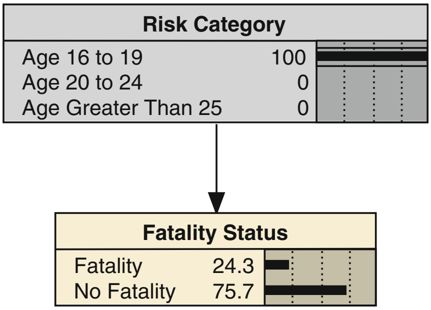

Fig. 9.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 9.3, not only can an analyst determine the probability of a fatality effecting an insured in Age 16–19, P(Age 16–19| Fatality) = 47.1% (Fig. 9.2), but she or he can also determine the percentage effect on Fatality (No Fatality) given an Age 16–19 insurer, P(Fatality|Age 16–19) = 24.3%, which are two distinct probabilities. Invoking Age 20–24 and Age > 25 will similarly adjust these conditional probabilities

9.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Risk Category, to the cause event, Fatality Status. For example in referring to Fig. 9.3, not only can an analyst determine the probability of an fatality effecting an insured in Age 16–19, P(Age 16–19|Fatality) = 47.1% (Fig. 9.2), but she or he can also determine the percentage effect on Fatality (No Fatality) given an Age 16–19 insurer, P(Fatality|Age 16–19) = 24.3%, which are two distinct probabilities. Invoking Age 20–24 or Age > 25 will similarly adjust these conditional probabilities.

Chapter 10
Acts of Terrorism (AOT) Example

10.1 Scenario

In this scenario, the U.S. dollar (USD) and currencies that are pegged to the USD continually appreciated and depreciates based on global economic conditions. The USD pegged currencies are experiencing depreciation across a range of economic effects due to whipsaw actions that generate 1,000 pip devaluation movements over short periods of time. The Office of the Controller of the Currency's (OCC) concern is that these currencies will continue to depreciate and exponentially increase the opportunity lost cost of these currencies in a global market place. Their research question is to determine the proportions of country currency depreciation rates given future economic effects. Maintaining optimal currency valuations with the minimal amount of economic costs would be a benefit to the countries of concern. OCC analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

Step 1: Identify a population of interest.

The population consists of AOTs against Americans in France, Germany, and Greece, historically.

The universe would consist of all AOT this country could have committed, and a subset would be the number of attempts made against the U.S. that were either successful or unsuccessful. To narrow down our universe, our sample population consists of the countries of France, Germany, and Greece.

10.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of countries of interest that have experienced AOT where Americans are present.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_10,

Table 10.1 Frequency counts

	Fatality status			
Country	Fatality	Injured	No harm	Total
France	34	68	590	692[a]
Germany	12	23	650	685
Greece	21	89	495	605
Total	67[b]	180	1,735	1,982

Note: These values represent country frequency counts for each of the Fatality Status elements
[a]692 = 34 + 68 + 590
[b]67 = 34 + 12 + 21

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The three disjoint elements are: "France," "Germany," and "Greece" from the element, "Country."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, Department of State (DOS) has reported AOT of 21.0% in France, 59.0% in Germany, and 20.0% in Greece from these respective embassies.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that can to cause harm to an American following an AOT. The disjoint event becomes "AOT." The analyst will slice through AOT by identifying the effects following AOT that result in a "Fatality," "Injured," or "No Harm," which become the sub-events for this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws from France, Germany, and Greece from an AOT database. The sampling process starts with a single random draw and selection of an element from Country and then from Fatality Status and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 10.1, Frequency Counts, for 692 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 10.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 10.2 across Country and Fatality Status. To compute marginal probabilities, the analyst then sums the joint

Table 10.2 Relative frequency/likelihood/conditional probabilities

	Fatality status			
Country	Fatality (%)	Injured (%)	No harm (%)	Total (%)
France	4.9[a]	9.8	85.3	100.0
Germany	1.8	3.4	94.9	100.0
Greece	3.5	14.7	81.8	100.0
Total	4.9	9.8	85.3	100.0
Conditional Probabilities				
Total	3.4[b]	9.1	87.5	100.0

Note: These values represent country relative frequencies/likelihood/conditional probabilities for each fatality status category that the analyst calculated using count data reported in Table 10.1

[a]4.9% = 34/692 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Transistor Quality and across Supplier using the frequency counts from Table 10.1

[b]3.4% = 67/1,982 × 100

Table 10.3 Joint and marginal probabilities

	Fatality status			
Country	Fatality (%)	Injured (%)	No harm (%)	Marginal probabilities (%)
France	1.0[a]	2.1	17.9	21.0[b]
Germany	1.0	2.0	56.0	59.0
Greece	0.7	2.9	16.4	80.0
Marginal Probabilities				
Total	2.8[c]	7.0	90.3	100.0

Notes: These values represent the joint probabilities for each Country and Fatality Status elements that the analyst calculated using prior probabilities and probabilities reported in Table 10.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Fatality Status and Country are dependent as evaluated by P(Fatality ∩ France) ≠ P(Fatality) × P(France), 1.3% ≠ 2.8% × 21.0% = 0.6%

[a]1.0% = 21.0% × 4.9%

[b]21.0% = 1.0% + 2.1% + 17.9%

[c]2.8% = 1.0% + 1.0% + 0.7%

probabilities down the elements of Country and Fatality Status, which totals 100.0%. The analyst then reports these in Table 10.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 10.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Fatality Status by dividing them individually by their respective marginal probabilities and then reports these in Table 10.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 10.1.

Table 10.4 Posterior probabilities

	Fatality status		
Country	Fatality (%)	Injured (%)	No harm (%)
France	37.4[a]	29.5	19.8
Germany	37.5	28.4	62.0
Greece	25.2	42.1	18.1
Total	100.0[b]	100.0	100.0

Note: This represents the posterior probabilities of the elements of Fatality Status. The analyst calculated them using the joint and marginal probabilities reported in Table 10.3

[a]37.4% = 1.0%/2.8%

[b]100.0% = 37.4% + 37.5% + 25.2%

Acts of Terrorism Example						
Country		Fatality Status	Likelihood	Joint	Marginal	Posterior
France	21.0%	Fatality	4.9%	1.0%	2.8%	37.4%
		Injured	9.8%	2.1%	7.0%	29.5%
		No Harm	85.3%	17.9%	90.3%	19.8%
Germany	59.0%	Fatality	1.8%	1.0%		37.5%
		Injured	3.4%	2.0%		28.4%
		No Harm	94.9%	56.0%		62.0%
Greece	20.0%	Fatality	3.5%	0.7%		25.2%
		Injured	14.7%	2.9%		42.1%
		No Harm	81.8%	16.4%		18.1%

Fig. 10.1 Tree diagram for the AOTs example. From Tables 10.1, 10.2, 10.3 and 10.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 10.2.

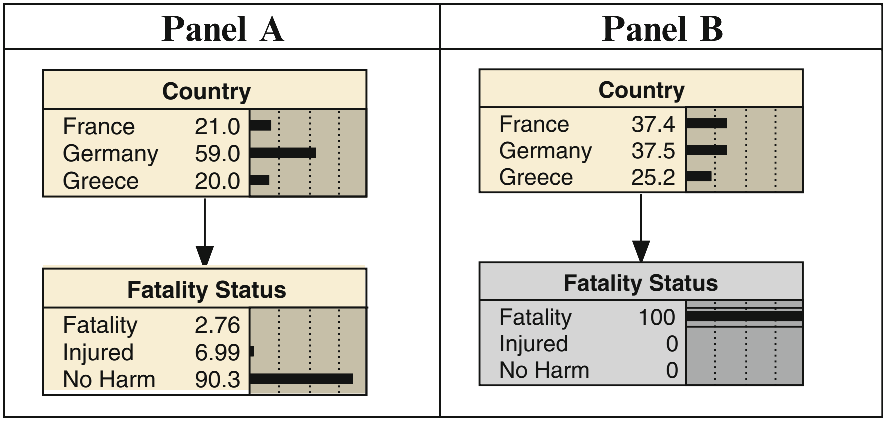

Fig. 10.2 This represents the Netica replication of this AOT example. *Panel A* represents the subjective or prior probabilities (Country Node) and the conditional or marginal joint probabilities (Fatality Status Node) (Table 10.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Fatality (Table 10.4). The analyst uses the raw data inputs into the Netica un-normalized input table

10.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

10.3.1 Posterior Probabilities

Given a fully saturated model containing all available information, an analyst can evaluate conditional probability changes going from the cause event, Fatality Status, to the effect event, Country. After conducting this experiment and priming the BBN with all available information, following the next AOT when an American citizen becomes a fatality, an analyst could conclude there is a 37.4% chance that this occurred in France, a 37.5% chance in Germany, and a 25.2% chance in Greece. An analyst can obtain similar revised probabilities for Injured and No Harm outcomes. Other events that could contribute to updating the posterior probabilities of this BBN could include unrest, economic embargos, and revenge attacks.

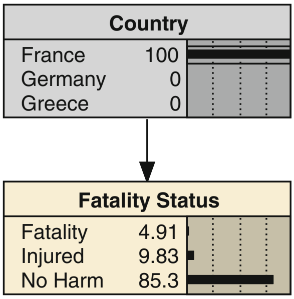

Fig. 10.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example, in referring to Fig. 10.3, not only can an analyst determine the probability of a fatality effecting an American in France, P(France| Fatality) = 37.4% (Fig. 10.2), but she or he can also determine the percentage effect on Fatality (Injured and No Harm) given an AOT in France, P(Fatality|France) = 4.91%, which are two distinct probabilities. Invoking Germany and Greece will similarly adjust these conditional probabilities

10.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Country, to the cause event, Fatality Status. For example in referring to Fig. 10.3, not only can an analyst determine the probability of an fatality effecting an American in France, P(France| Fatality) = 37.4% (Fig. 10.2), but she or he can also determine the percentage effect on Fatality (Injured and No Harm) given an AOT in France, P(Fatality| France) = 4.91%, which are two distinct probabilities. Invoking Germany or Greece will similarly adjust these conditional probabilities.

Acknowledgments I would like to thank Major Tonya R. Tatum for her contributions to this chapter. Major Tatum is an Operations Research/Systems Analyst with the Mission and Recruiter Requirements Division, Assistant Chief of Staff, G2, U.S. Army Recruiting Command.

Chapter 11
Currency Wars Example*

11.1 Scenario

In this scenario, the U.S. dollar (USD) and currencies that are pegged to the USD continually appreciated and depreciates based on global economic conditions. The USD pegged currencies are experiencing depreciation across a range of economic effects due to whipsaw actions that generate 1,000 pip devaluation movements over short periods of time. The Office of the Controller of the Currency's (OCC) concern is that these currencies will continue to depreciate and exponentially increase the opportunity lost cost of these currencies in a global market place. Their research question is to determine the proportions of country currency depreciation rates given economic effects. Maintaining optimal currency valuations with the minimal amount of economic costs would be a benefit to the countries of concern. The OCC analyst will evaluate this scenario using a Bayesian Belief Network (BBN).

11.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of weekly price movements greater than or equal to 1,000 pips during a 10-year period.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. There are two disjoint elements are the following currency pairs: USD/CHF, USD/JPY, and EUR/USD.

*The idea of a currency war comes from Rickard (Rickard 2012). I obtained date from the OANDA website (OANDA 2012).

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_11,

Table 11.1 Frequency counts

Currency pair	Economic effects			Total
	Natural disaster	Assassination	National bank intervention	
USD/CAD	4	3	8	15[a]
EUR/USD	2	5	8	15
USD/JPY	8	2	7	17
Total	14[b]	10	23	47

Note: These values represent currency pair frequency counts for each of the Economic Effects elements

[a]25 = 4 + 3 + 8

[b]14 = 4 + 2 + 8

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, this basket of currencies has experiences the following pip movement mix: (1) USD/CAD, 22.0%, (2) EUR/USD, 36.0%, and (3) USD/JPY, 42.0%.[1]

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for events can cause pip whipsaw movements in excess of 1,000 pips in a week. The disjoint event is "Economic Event." The analyst would slice through Economic Event by identifying the effects following the following events: (1) Natural Disasters, (2) Assassinations, and (3) National Bank Interventions, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of currency pairs: (1) USD/CAD, (2) EUR/USD, and (3) USD/JPY from a database. The sampling process starts with a single random draw and selection of an element from Currency Pair and then from Economic Event and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 11.1, Frequency Counts, for 15 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 11.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 11.2 across Currency

[1] I obtained currency data from: http://www.oanda.com/.

Table 11.2 Relative frequency/likelihood/conditional probabilities

	Economic event			
Currency pair	Natural disaster (%)	Assassination (%)	National bank intervention (%)	Total (%)
USD/CAD	26.7[a]	20.0	53.3	100.0
EUR/USD	13.3	33.3	53.3	100.0
USD/JPY	47.1	11.8	41.2	100.0
Conditional probabilities				
Total	29.8[b]	21.3	48.9	100.0

Note: These values represent currency pairs' relative frequencies/likelihood/conditional probabilities for each economic event that the analyst calculated using count data reported in Table 11.1

[a]26.7% = 4/15 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Transistor Quality and across Supplier using the frequency counts from Table 11.1

[b]29.8% = 14/47 × 100

Table 11.3 Joint and marginal probabilities

	Economic event			
Currency pair	Natural disaster (%)	Assassination (%)	National bank intervention (%)	Marginal probabilities (%)
USD/CAD	5.9[a]	4.4	11.7	22.0[b]
EUR/USD	4.8	12.0	19.2	36.0
USD/JPY	19.8	4.9	17.3	58.0
Marginal probabilities				
Total	30.4[c]	21.3	48.2	100.0

Notes: These values represent the joint probabilities for each Currency Pair and Economic Event elements that the analyst calculated using prior probabilities and probabilities reported in Table 11.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Economic Event and Currency Pair are dependent as evaluated by P(Natural Disaster ∩ USD/CAD) ≠ P(Natural Disaster) × P(USD/CAD), 5.9% ≠ 30.4% × 22.0% = 6.7%

[a]5.9% = 22.0% × 22.7%

[b]22.0% = 5.9% + 4.40% + 11.7%

[c]30.4% = 5.9% + 4.8% + 19.8%

Pair and Economic Event. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Currency Pair and Economic Events, which totals 100.0%. The analyst then reports these in Table 11.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 11.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Economic Event by dividing them individually by

Table 11.4 Posterior probabilities

	Economic event		
Currency pair	Natural disaster (%)	Assassination (%)	National bank intervention (%)
USD/CAD	19.3[a]	20.6	24.3
EUR/USD	15.8	56.2	39.8
USD/JPY	64.9	23.2	35.9
Total	100.0[b]	100.0	100.0

Note: This represents the posterior probabilities of the elements of Economic Event. The analyst calculated them using the joint and marginal probabilities reported in Table 11.3
[a]19.3% = 5.9%/30.4%
[b]100.0% = 19.3% + 15.8% + 64.9%

Currency Wars Example

Currency Pair		Economic Event	Likelihood	Joint	Marginal	Posterior
USD/CAD	22.0%	Natural Disaster	26.7%	5.9%	30.4%	19.3%
		Assasination	20.0%	4.4%	21.3%	20.6%
		National Bank Intervention	53.3%	11.7%	48.2%	24.3%
EUR/USD	36.0%	Natural Disaster	13.3%	4.8%		15.8%
		Assasination	33.3%	12.0%		56.2%
		National Bank Intervention	53.3%	19.2%		39.8%
USD/JPY	42.0%	Natural Disaster	47.1%	19.8%		64.9%
		Assasination	11.8%	4.9%		23.2%
		National Bank Intervention	41.2%	17.3%		35.9%

Fig. 11.1 Tree diagram for the Currency Wars example. From Tables 11.1, 11.2, 11.3 and11.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

their respective marginal probabilities and then reports these in Table 11.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 11.1.

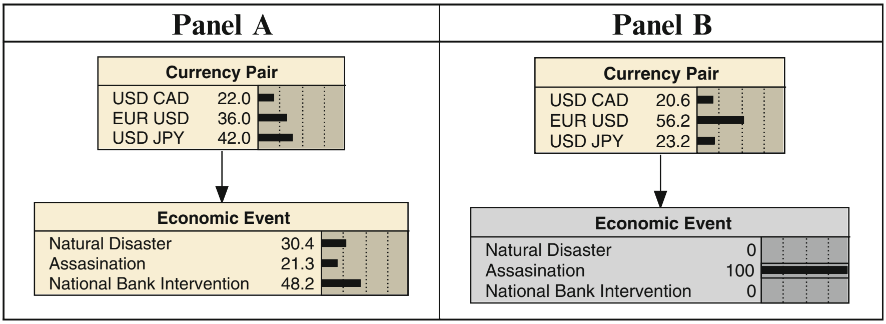

Fig. 11.2 This represents the Netica replication of this Currency Wars example. *Panel A* represents the subjective or prior probabilities (Currency Pair Node) and the conditional or marginal joint probabilities (Economic Event Node) (Table 11.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Assassination (Table 11.4). The analyst uses the raw data inputs into the Netica un-normalized input table

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 11.2.

11.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

11.3.1 Posterior Probabilities

Here, an analyst can evaluate conditional probability changes going from the effect event, Currency Pair, to the cause event, Economic Event. After conducting this experiment and priming the BBN with all available information, if the next 1,000 pip movement is caused by an assassination, then she or he could conclude there is a 26.1% chance that it will affect the USD/CAD pair, a 56.2% the EUR/USD pair, and a 23.2% the USD/JPY pair. Other events that could contribute to the updating of the posterior probabilities of this BBN include county inflation rates and the release of new economic data.

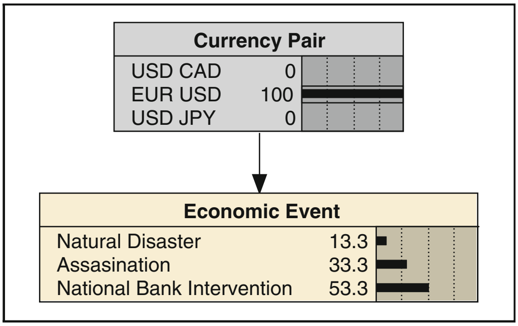

Fig. 11.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 11.3, not only can an analyst determine the probability of an assassination effecting a 1,000 pip currency movement, P(EUR/USD|Assassination) = 20.6% (Fig. 11.2), but she or he can also determine the percentage effect on Assassination (Natural Disaster and National Bank Intervention) given a 1,000 pip movement in the EUR/USD pair, P(Assassination|EUR/USD) = 33.3%, which are two distinct probabilities. Invoking USD/CAD and USD/JPY will similarly adjust these conditional probabilities

11.3.2 Inverse Probabilities

Using inverse probabilities, an analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Currency Pair, to the cause event, Economic Event. For example in referring to Fig. 11.3, not only can an analyst determine the probability of an assassination effecting a 1,000 pip currency movement, P(EUR/USD|Assassination) = 20.6% (Fig. 11.2), but she or he can also determine the percentage effect on Assassination (Natural Disaster and National Bank Intervention) given a 1,000 pip movement in the EUR/USD pair, P(Assassination|EUR/USD) = 33.3%, which are two distinct probabilities. Invoking USD/CAD and USD/JPY will similarly adjust these conditional probabilities.

References

OANDA (2012). *OANDA forex trading*. http://www.oanda.com/. Accessed 1 May 2012.
Rickard, J. (2012). *Currency wars: The making of the next global crisis*. New York: Penguin.

Chapter 12
College Entrance Exams Example

12.1 Scenario

In this scenario, a university obtains freshmen students based on multiple with American College Testing (ACT) score levels. The admissions department is experiencing an above-average number of freshmen dropouts across each level of ACT scores. Their concern is that these freshmen will continue to drop out and exponentially increase the opportunity lost cost of these students and also send a signal to the accreditation authority of possible creditability issues. Their research question is to determine the proportions of freshman maturation given the levels of ACT scores. Obtaining quality freshman with the minimal amount of costs would be a benefit to the university. A university analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

12.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of freshmen whom the university has historically accepted.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. There are two disjoint elements are "Graduate" and "Non-Graduate" freshmen from the element, "Freshman Status."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the 85.0% of Freshman who have been accepted and started have graduated in 4 years.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that can determine which entering freshman would actually graduate in 4 years. The analyst disjoint event is "ACT Scores." The analyst will slice through

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_12,

Table 12.1 Frequency counts

	ACT scores			
Freshman status	Level 1	Level 2	Level 3	Total
Graduate	369	52	38	459[a]
Non graduate	83	58	26	167
Total	452[b]	110	64	626

Note: These values represent transistor freshman status frequency counts for each of the ACT Scores elements
[a]459 = 369 + 52 + 38
[b]452 = 369 + 83

Table 12.2 Relative frequency/likelihood/conditional probabilities

	ACT scores			
Freshman status	Level 1 (%)	Level 2 (%)	Level 3 (%)	Total (%)
Graduate	80.4[a]	11.3	8.3	100.0
Non graduate	49.7	34.7	15.6	100.0
Conditional probabilities				
Total	72.2[b]	34.7	10.2	100.0

Note: These values represent freshman status relative frequencies/likelihood/conditional probabilities for each ACT Score that the analyst calculated using count data reported in Table 12.1
[a]80.4% = 369/459 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Transistor Quality and across Supplier using the frequency counts from Table 12.1
[b]72.2% = 452/426 × 100

ACT Scores by identifying the effects from ACT score ranges of (1) Level 1, (2) Level 2, and (3) Level 3, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of Graduates and Non-Graduates from a database. The sampling process starts with a single random draw and selection of an element from Freshman Status and then from ACT Scores and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 12.1, Frequency Counts, for 459 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Transistor Quality event. To determine these percentages, the analyst calculates probabilities across the sliced events of Supplier and then reports these results in Table 12.2, Relative Frequency/Likelihood/Conditional Probabilities.

Table 12.3 Joint and marginal probabilities

	ACT scores			
Freshman status	Level 1 (%)	Level 2 (%)	Level 3 (%)	Marginal probabilities (%)
Graduate	68.3[a]	9.6	7.0	85.0[b]
Not Graduate	7.5	5.2	2.3	15.0
Marginal Probabilities				
Total	75.8[c]	14.8	9.4	100.0

Notes: These values represent the joint probabilities for each Freshman Status and ACT Score elements that the analyst calculated using prior probabilities and probabilities reported in Table 12.2, Relative Frequency/Likelihood/Conditional Probabilities. Events ACT Scores and Freshman Status are dependent as evaluated by P(Level 1∩ Graduate) $\neq$ P(Level 1) $\times$ P(Graduate), 68.3% $\neq$ 75.8% $\times$ 85.0% = 64.4%

[a]68.3% = 85.0% $\times$ 80.4%

[b]85.0% = 68.3% + 9.6% + 7.0%

[c]75.8% = 68.3% + 74.6%

Table 12.4 Posterior probabilities

	ACT scores		
Freshman status	Level 1 (%)	Level 2 (%)	Level 3 (%)
Graduate	90.2[a]	64.9	75.1
Not graduate	9.8	35.1	24.9
Total	100.0[b]	100.0	100.0

Note: This represents the posterior probabilities of the elements of ACT Scores. The analyst calculated them using the joint and marginal probabilities reported in Table 12.3

[a]90.2% = 68.3% / 75.8%

[b]100.0% = 90.2% + 9.8%

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 12.2 across Freshman Status and ACT Scores. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Freshman Status and ACT Scores, which totals 100.0%. The analyst then reports these in Table 12.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 12.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in ACT Scores by dividing them individually by their respective marginal probabilities and then reports these in Table 12.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 12.1.

Step 11: Run a netica peplication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 12.2.

College Entrance Exams Example

Graduate		Freshman Status	Likelihood	Joint	Marginal	Posterior
Graduate	85.0%	Level 1	80.4%	68.3%	75.8%	90.2%
		Level 2	11.3%	9.6%	14.8%	64.9%
		Level 3	8.3%	7.0%	9.4%	75.1%
Not Graduate	15.0%	Level 1	49.7%	7.5%		9.8%
		Level 2	34.7%	5.2%		35.1%
		Level 3	15.6%	2.3%		24.9%

Fig. 12.1 Tree diagram for the College Entrance Exams example. From Tables 12.1 through 12.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

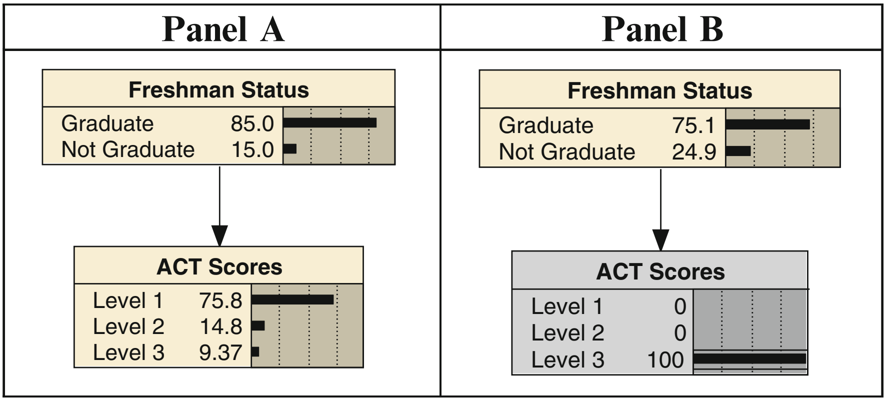

Fig. 12.2 This represents the Netica replication of this College Entrance Exams example. *Panel A* represents the subjective or prior probabilities (Freshman Status Node) and the conditional or marginal joint probabilities (ACT Scores Node) (Table 12.3). *Panel B* represents the revised prior or posterior probabilities when the analyst invoked the element Level 3 (Table 12.4). The analyst uses the raw data inputs into the Netica un-normalized input table

12.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

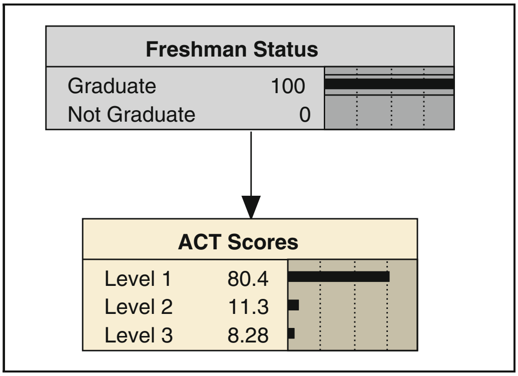

Fig. 12.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 12.3, not only can an analyst determine the probability of level 3 students graduating in 4 years, P (Graduate|Level 3) = 75.1% (Fig. 12.2), but she or he can also determine the percentage effect of level 3 (level 1 and level 2) students graduating in 4 years, P(Level 3|Graduate) = 80.4%, which are two distinct probabilities. Invoking not graduate will similarly adjust these conditional probabilities

12.3.1 Posterior Probabilities

Given a fully saturated model containing all available information, an analyst can evaluate conditional probability changes going from the cause event, ACT Scores, to the effect event, Freshman Status. After conducting this experiment and priming the BBN with all available information, following the next freshman admission of a student who scores a Level 3 on the ACT, then an analyst can conclude that she or he has a 75.1% chance of and graduating and a 24.9% of not graduating in 4 years. An analyst can obtain revised conditional probabilities when invoking Level 1 and Level 2 similarly. Other events that could contribute to updating the posterior probabilities of this BBN could include a student's desire as measured by a psychological measurement tool, parental support and income level, and possibly sex.

12.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Freshman Status, to the cause event, ACT Scores. For example in referring to Fig. 12.3, not only can an analyst determine the probability of Level 3 students graduating in four years, P(Graduate|Level 3) = 75.1% (Fig. 12.2), but an analyst can also determine the percentage effect of Level 3 (Level 1 and Level 2) students graduating in four years, P(Level 3|Graduate) = 80.4%, which are two distinct probabilities. Invoking Not Graduate will similarly adjust these conditional probabilities. This represents the Level 3 portion of all graduating students, Levels 1, 2, and 3.

Chapter 13
Special Forces Assessment and Selection (SFAS) One-Stage Example*

13.1 Scenario

In this scenario, the U.S. Army Special Forces Command's (USASFC) Special Forces Assessment and Selection (SFAS) course obtain Soldiers from the ranks of the Army. The SFAS has experienced an elevated level of attrition rates of Soldiers they are receiving from the current recruiting of enlisted and Officer Soldiers.[1] Their concern is that these high attrition rates will stop the Special Forces community from being fully mission capable, according to regulatory requirements and increase the cost of recruiting and assessing future Soldiers. Their research question is to determine the proportions of not selected and selected Soldiers given the Soldier is enlisted or officer and their physical fitness levels. Selecting the right Soldier with the minimal amount of costs would be a benefit to the U.S. Army's recruiting program. An Army Research Institute analyst will evaluate this scenario using a One-Stage Bayesian Belief Network (BBN).

13.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of Soldiers who were accepted and entered into SFAS.

*See the U.S. Army's Special Forces website: http://www.sorbrecruiting.com/. Last accessed: 11/6/2012.

[1] The SFAS course is a 3-week evaluation of enlisted and Officer Soldiers' physical, mental, and psychological capabilities to determine if they would fit the ranks of special operations Soldiers. Those Soldiers accepted through SFAS will attend either the Officer or Enlisted Special Forces Qualification Course (SFQC) for final selection to earn the Green Beret. SFAS is only a gateway to the SFQC.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_13,

Table 13.1 Frequency counts

	Status		
Graduate	Enlisted	Officer	Total
Selected	663	52	715[a]
Not selected	263	22	285
Total	926[b]	74	1,000

Note: These values represent selection frequency counts for each of the Status elements
[a]715 = 663 + 52
[b]926 = 663 + 263

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population. The two disjoint elements are "Selected" and "Not Selected" Soldiers from the element, "Graduate."

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the selection rate has been 30.0% for all Soldiers who were accepted and entered into SFAS.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements. In this example, the analyst is looking for an event with an outcome that can identify which Soldier will not be selected. The disjoint event is "Status." The analyst will slice through Status by identifying the effects following the entrance of these Soldiers into SFAS from the enlisted and officer ranks, which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of Selected and Not Selected Soldiers from a database. The sampling process starts with a single random draw and selection of an element from Graduate and then from Status and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count data for further analysis. The analyst has reports these results in Table 13.1, Frequency Counts, for 715 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies). The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Graduate event. To determine these percentages, the analyst calculates probabilities across the sliced events of Status and then reports these results in Table 13.2, Relative Frequency/Likelihood/Conditional Probabilities.

Step 8: Determine joint and marginal probabilities. To compute joint probabilities, the analyst multiplies the likelihood probabilities in Table 13.2 across Graduate and Status. To compute marginal probabilities, the analyst then sums the joint

Table 13.2 Relative frequency/likelihood/conditional probabilities

	Status		
Graduate	Enlisted (%)	Officer (%)	Total (%)
Selected	92.7[a]	7.3	100.0
Not selected	92.3	7.7	100.0
Conditional probabilities			
Total	92.6[b]	7.4	100.0

Note: These values represent selection relative frequencies/likelihood/conditional probabilities for each Soldier status that the analyst calculated using count data reported in Table 13.1
[a]92.7% = 663/715 × 100. The analyst computed the marginal probabilities by dividing the total frequency counts down Graduate and across Status using the frequency counts from Table 13.1
[b]92.6% = 926/1,000 × 100

Table 13.3 Joint and marginal probabilities

	Status		
Graduate	Enlisted (%)	Officer (%)	Marginal probabilities (%)
Selected	27.8[a]	2.2	30.0[c]
Not selected	64.6	5.4	70.0
Marginal probabilities			
Total	92.4[b]	7.6	100.0

Notes: These values represent the joint probabilities for each Graduate and Status elements that the analyst calculated using prior probabilities and probabilities reported in Table 13.2, Relative Frequency/Likelihood/Conditional Probabilities. Events Status and Graduate are dependent as evaluated by P(Enlisted ∩ Selected) ≠ P(Enlisted) × P(Selected), 27.8%≠ 92.4% × 30.0% = 27.7%
[a]27.8% = 30.0% × 92.7%
[b]92.4% = 27.8% + 64.6%
[c]30.0% = 27.8% + 2.2%

probabilities down the elements of Graduate and Status, which totals 100.0%. The analyst then reports these in Table 13.3, Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities. To compute posterior probabilities, the analyst divides the joint probabilities in Table 13.3 by their respective marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Status by dividing them individually by their respective marginal probabilities and then reports these in Table 13.4 as posterior probabilities.

Step 10: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 13.1.

Step 11: Run a netica replication. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Fig. 13.2.

Table 13.4 Posterior probability table

	Status	
Graduate	Enlisted (%)	Officer (%)
Selected	30.1[a]	28.8
Not selected	69.9	71.2
Total	100.0[b]	100.0

Note: This represents the posterior probabilities of the elements of Status. The analyst calculated them using the joint and marginal probabilities reported in Table 13.3

[a]30.1% = 27.8%/92.4%

[b]100.0% = 30.1% + 69.9%

Special Forces Assessment and Selection Example

Graduate		Status	Likelihood	Joint	Marginal	Posterior
Selected	30.0%	Enlisted	92.7%	27.8%	92.4%	30.1%
		Officer	7.3%	2.2%	7.6%	28.8%
Not Selected	70.0%	Enlisted	92.3%	64.6%		69.9%
		Officer	7.7%	5.4%		71.2%

Fig. 13.1 Tree diagram for the SFAS example. From Tables 13.1, 13.2, 13.3 and 13.4, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this One-Stage BBN Model

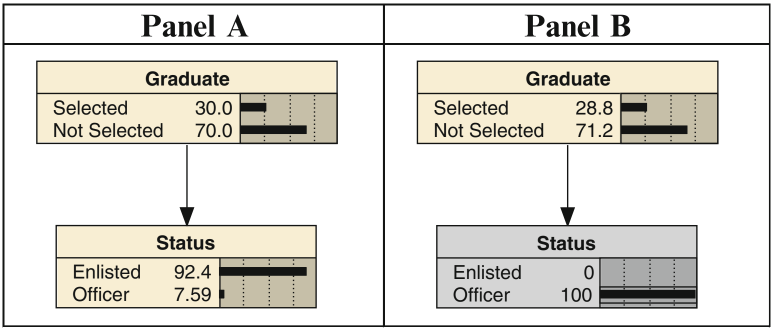

Fig. 13.2 This represents the Netica replication of this SFAS stage one BBN example. *Panel A* represents the subjective or prior probabilities (Graduate Node) and the conditional or marginal joint probabilities (Status Node) (Table 13.3). *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the element Officer (Table 13.4). The analyst uses the raw data inputs into the Netica un-normalized input table

13.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

13.3.1 Posterior Probabilities

Given a fully saturated model containing all available information, an analyst can evaluate conditional probability changes going from the cause event, Status, to the effect event, Graduate. After conducting this experiment and priming the BBN with all available information, following the next officer completing SFAS, an analyst could conclude there is a 28.8% chance that he will and 71.2% that he will not be selected to attend the SFQC. An analyst can obtain similar revised probabilities for enlisted Soldier outcomes. Other events that could contribute to updating the posterior probabilities of this BBN could include a Soldiers physical training level.[2]

13.3.2 Inverse Probabilities

Using inverse probabilities, an analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Graduate, to the cause event, Status. For example in referring to Fig. 13.3, not only can an analyst determine the probability of officers being selected at SFAS, P(Selected|Officer) = 28.8% (Fig. 13.2), but she or he can also determine the percentage effect of selected officers (and enlisted), P(Officer|Selected) = 7.27% (Fig. 13.3), which are two distinct probabilities. Invoking Not Selected will similarly adjust these conditional probabilities. This represents the Selected portion of all selected Soldiers, enlisted and officer.[3]

[2] I will include this variable during my evaluation of the SFAS Two Stage Model in Chap. 14.

[3] Due to very low selection rates and the smaller proportion of officers attending SFAS, this percentage is also low.

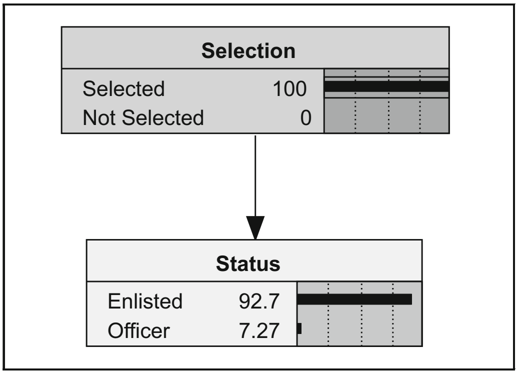

Fig. 13.3 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example, in referring to Fig. 13.3, not only can an analyst determine the probability of officers being selected at SFAS, P(Selected|Officer) = 28.8% (Fig. 13.2), but she or he can also determine the percentage effect of selected officers (and enlisted), P(Officer|Selected) = 7.27%, which are two distinct probabilities. Invoking Enlisted will similarly adjust these conditional probabilities

Reference

Special Forces (2012). http://www.sorbrecruiting.com/. Last accessed: 11/6/2012.

Chapter 14
Special Forces Assessment and Selection (SFAS) Two-Stage Example*

14.1 Scenario

In this scenario, the U.S. Army Special Forces Command's (USASFC) Special Forces Assessment and Selection (SFAS) course obtains Soldiers from the ranks of the Army. The SFAS has experienced an elevated level of attrition rates of Soldiers they are receiving from the current recruiting of enlisted and Officer Soldiers.[1] Their concern is that these high attrition rates will stop the Special Forces community from being fully mission capable, according to regulatory requirements and increase the cost of recruiting and assessing future Soldiers. Their research question is to determine the proportions of not selected and selected Soldiers given the Soldier is enlisted or officer and their physical fitness levels. Selecting the right Soldier with the minimal amount of costs would be a benefit to the U.S. Army's recruiting program. An Army Research Institute analyst will evaluate this scenario using a Two-Stage Bayesian Belief Network (BBN).

14.2 Experimental Protocol

Step 1: Identify a population of interest. The population consists of the total number of Soldiers who were accepted and entered into SFAS.

Step 2: Slice through this population and identify at a minimum two mutually exclusive or disjoint (unconditional) events, which are the subsets of our population.

* See the U.S. Army's Special Forces website: http://www.sorbrecruiting.com/. Last accessed: 11/6/2012.

[1] See the U.S. Army's Special Forces website: http://www.sorbrecruiting.com/. Last accessed: 11/6/2012. The SFAS course is a 3-week evaluation of enlisted and Officer Soldiers' physical, mental, and psychological capabilities to determine if they would fit the ranks of special operations Soldiers. Those Soldiers accepted through SFAS will attend either the Officer or Enlisted Special Forces Qualification Course (SFQC) for final selection to earn the Green Beret. SFAS is only a gateway to the SFQC.

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4_14,

In this Two Event scenario, the two disjoint elements are "Selected" and "Not Selected" Soldiers from the element, "Graduate" and "Above," "Extreme," and "Average" from the element, Physical Training (PT).

Step 3: Determine prior (a priori) or unconditional probabilities. Historically, the selection rate has been 28.5% for all Soldiers who were accepted and entered into SFAS.

Step 4: Identify the conditional event and its subset of mutually exclusive or disjoint (unconditional) elements.

Stage 1. In this example, the analyst is looking for an event than can cause a Soldier to not be selected following the attendance to SFAS. The disjoint event becomes "Rank." The analyst will slice through Status by identifying the effects following the entrance of these Soldiers into SFAS from the enlisted and officer ranks, which become the elements of this event.

Stage 2. In this example, the analyst is looking for an event than can cause a Soldier to not be selected following the attendance to SFAS. The disjoint event becomes "PT." The analyst will slice through PT by identifying the effects following the entrance of these Soldiers into SFAS from the enlisted and officer ranks and their PT status of "Above," "Extreme," or "Average," which become the elements of this event.

Step 5: Conduct the random experiment. The analyst performs this experiment by making random draws of Selected and Not Selected Soldiers from a database. The sampling process starts with a single random draw and selection of an element from Graduate, Status, and then from PT and ends with the assignment of the draw results. The analyst will continue this process until she or he has obtained the desired sample size.

Step 6: Determine frequency counts. To record frequencies, the analyst reports count for further analysis for both nodes in Table 14.1, Frequency Counts-Stage 1 Status Node for 1,000 iterations; Table 14.2, Frequency and Adjusted Frequency Counts Stage 2 Status Node-Enlisted for 926 iterations; Table 14.3 Frequency and Adjusted Frequency Counts Stage 2 Status Node-Officer for 76 iterations; and finally sums these counts in Table 14.4, Stage 2 Status Node-Total Officer & Enlisted, again for 1,000 iterations.

Step 7: Determine likelihood/conditional probabilities (relative frequencies).

Stage 1. The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Graduate event. To determine these percentages, the analyst calculates probabilities across the sliced events of Status Node-Stage 1 and then reports these results in Table 14.5, Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1.

Stage 2. The analyst then computes relative frequencies/likelihood/conditional probabilities as conditional probabilities based on the subjective probabilities of the Graduate event for both Enlisted and Officer. To determine these percentages,

Table 14.1 Frequency counts-stage 1 status node

	Stage 1-status node-total		
Graduate	Enlisted	Officer	Total
Not selected	663	52	715[a]
Selected	263	22	285
Total	926[b]	74	1,000

Note: These values represent transistor quality frequency counts for each of the Stage 1-Status Node-Total elements
[a]715 = 663 + 52
[b]926 = 663 + 263

Table 14.2 Frequency and adjusted frequency counts stage 2 status node-enlisted

	Stage 2-PT node-enlisted			
Graduate	Above	Extreme	Average	Total
Not selected	237	234	192	663[a]
Selected	92	82	89	263
Total	329[b]	316	281	926

Note: These values represent transistor quality frequency counts for each of the Stage 2-PT Node-Enlisted elements
[a]663 = 237 + 234 + 192
[b]329 = 237 + 92

Table 14.3 Frequency and adjusted frequency counts stage 2 status node-officer

	Stage 1-PT node-officer			
Graduate	Above	Extreme	Average	Total
Not selected	18	16	18	52[a]
Selected	9	6	7	22
Total	27[b]	22	25	74

Note: These values represent transistor quality frequency counts for each of the Stage 1-PT Node-Officer elements
[a]52 = 18 + 16 + 18
[b]27 = 18 + 9

Table 14.4 Stage 2 status node-total officer and enlisted

	Stage 2-PT node-total			
Graduate	Above	Extreme	Average	Total
Not selected	255	250	210	715[a]
Selected	101	88	96	285
Total	356[b]	338	306	1,000

Note: These values represent transistor quality frequency counts for each of the Stage 2-PT Node-Total elements
[a]715 = 255 + 250 + 210
[b]356 = 255 + 101

Table 14.5 Relative frequency/likelihood/conditional probabilities-status node-stage 1

	Status node-stage 1		
Graduate	Enlisted (%)	Officer (%)	Total (%)
Not selected	92.7[a]	7.3	100.0
Selected	92.3	7.7	100.0
Conditional probabilities			
Total	92.6[b]	7.4	100.0

Note: These values represent transistor quality relative frequencies/likelihood/conditional probabilities for each Enlisted and Officer Soldier Status Node-Stage 1 that the analyst calculated using count data reported in Table 14.1 Frequency Counts-Stage 1 Status Node

[a]92.7% = 663/715 × 100. The analyst computed the conditional/marginal probabilities by dividing the total frequency counts down Graduate and across Status Node-Stage 1 using the frequency counts from Table 14.1

[b]92.6% = 926/1,000 × 100

Table 14.6 Relative frequency/likelihood/conditional probabilities-PT node-enlisted/officer-stage 2

	Stage 2-PT node-enlisted			
Graduate	Above (%)	Extreme (%)	Average (%)	Total (%)
Not selected	35.7[a]	35.3	29.0	100.0
Selected	35.0	31.2	33.8	100.0
	Stage 2-PT node-officer			
Graduate	Above	Extreme	Average	Total
Not selected	34.6[b]	30.8	34.6	100.0
Selected	40.9	27.3	31.8	100.0
	Conditional probabilities			
Total	35.6[c]	33.8	30.6	100.0

Note: (1) Stage 2-PT Node-Enlisted. These values represent transistor quality relative frequencies/likelihood/conditional probabilities for each PT level that the analyst calculated using count data reported in Table 14.2 Frequency and Adjusted Frequency Counts Stage 2 Status Node-Enlisted. (2) Stage 2-PT Node-Officer. These values represent transistor quality relative frequencies/likelihood/conditional probabilities for each PT level that the analyst calculated using count data reported in Table 14.3 Frequency and Adjusted Frequency Counts Stage 2 Status Node-Officer. (3) The analyst computed the conditional/marginal probabilities by dividing the total frequency counts down Graduate and across Stage 2-PT Node-Total using the frequency counts from Table 14.4 Stage 2 Status Node-Total Officer & Enlisted

[a]35.7% = 237/663 × 100

[b]34.6% = 18/52 × 100

[c]35.6% = 356/1,000 × 100

the analyst calculates probabilities across the sliced events of Stage 2-PT Node-Enlisted and Stage 2-PT Node-officer and then reports these results in Table 14.6, Relative Frequency/Likelihood/Conditional Probabilities-PT Node-Enlisted/Officer-Stage 2.

Step 8: Determine joint and marginal probabilities.

Stage 1. To compute joint probabilities, the analyst multiplies the prior probabilities by the respective probabilities in Table 14.5 Relative Frequency/

Table 14.7 Stage 1-Joint and marginal probabilities

	Status node-stage 1		
Graduate	Enlisted (%)	Officer (%)	Marginal probabilities (%)
Not selected	66.3[a]	5.2	71.5[c]
Selected	26.3	2.2	28.5
Marginal probabilities			
Total	92.6[b]	7.4	100.0

Notes: These values represent the joint probabilities for each Graduate and Status Node-Stage 1 elements that the analyst calculated using prior probabilities and count data reported in Table 14.5 Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1. Events Supplier and Transistor Quality are dependent as evaluated by P(Enlisted ∩ Not Selected) $\neq$ P(Enlisted) $\times$ P(Not Selected), 66.3% $\neq$ 92.6% $\times$ 71.5% = 66.2%

[a]66.3% = 35.7% $\times$ 71.5%. The analyst computed the Marginal Probabilities by summing down Status Node-Stage 1 and across Graduate elements

[b]92.6% = 31.2% + 8.1%

[c]71.5% = 66.3% + 5.2%

Likelihood/Conditional Probabilities-Status Node-Stage 1. To compute marginal probabilities, the analyst then sums the joint probabilities down the elements of Graduate and Status Node-Stage 1, which totals 100.0%. The analyst then reports these in Table 14.7, Stage 1-Joint and Marginal Probabilities.

Stage 2a. To compute joint probabilities for the Stage 2-PT Node-Enlisted, the analyst multiplies the prior probabilities by the respective probabilities in Table 14.5 Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1 by the probabilities in Table 14.6, Relative Frequency/Likelihood/Conditional Probabilities-PT Node-Enlisted/Officer-Stage 2. To compute conditional (marginal) probabilities, the analyst then sums the joint probabilities down the elements of Graduate and Stage 2-PT Node-Enlisted, which totals 100.0%. The analyst then reports these in Table 14.8 as joint, marginal, and conditional probabilities.

Stage 2b. To compute joint probabilities for the Stage 2-PT Node-Officer, the analyst multiplies the prior probabilities by the probabilities in Table 14.5 Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1 by the respective probabilities in Table 14.6, Relative Frequency/Likelihood/Conditional Probabilities-PT Node-Enlisted/Officer-Stage 2. To compute conditional (marginal) probabilities, the analyst then sums the joint probabilities down the elements of Graduate and Stage 2-PT Node-Officer, which totals 100.0%. The analyst then reports these in Table 14.8, Stage 2-Joint and Marginal Probabilities.

Step 9: Determine posterior probabilities.

Stage 1. To compute posterior probabilities, the analyst divides the joint probabilities in Table 14.7 Stage 1-Joint and Marginal Probabilities by their respective conditional/marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Status Node-Stage 1 by dividing them individually by their respective conditional/marginal probabilities and then reports these in Table 14.9 as posterior probabilities.

Table 14.8 Stage 2-Joint and marginal probabilities

	Stage 2-PT node-enlisted			
Graduate	Above (%)	Extreme (%)	Average (%)	Marginal probabilities (%)
Enlisted	23.7[a]	23.4	19.2	66.3[c]
Officer	9.2	8.2	8.9	26.3
	Marginal probabilities			
Sub-total	32.9[b]	31.6	28.1	92.6
	Stage 2-PT node-officer			
Graduate	Above (%)	Extreme (%)	Average (%)	Marginal probabilities (%)
Not selected	1.8[d]	1.6	1.8	5.2[f]
Selected	0.9	0.6	0.7	2.2
	Sub-marginal probabilities			
Sub-Total	2.7[e]	2.2	2.5	7.4
Marginal probabilities				
Total	35.6[g]	33.8	30.6	100.0

Notes: (1) Stage 2-PT Node-Enlisted. These values represent the joint probabilities for each Graduate and Stage 2-PT Node-Enlisted elements that the analyst calculated using prior probabilities, probabilities in Table 14.5 Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1, and probabilities in Table 14.6 Relative Frequency/Likelihood/Conditional Probabilities-PT Node-Enlisted/Officer-Stage 2. (2) Marginal Probabilities are computed by summing down each Stage 2-PT Node-Enlisted and across each Graduate element. (3) Stage 2-PT Node-Officer. These values represent the joint probabilities for each Graduate and Stage 2-PT Node-Officer elements that the analyst calculated using prior probabilities, probabilities in Table 14.5 Relative Frequency/Likelihood/Conditional Probabilities-Status Node-Stage 1, and probabilities in Table 14.6 Relative Frequency/Likelihood/Conditional Probabilities-PT Node-Enlisted/Officer-Stage 2. (4) Marginal Probabilities are computed by summing down each Stage 2-PT Node-Enlisted and across each Graduate element

[a] $23.7\% = 35.7\% \times 92.7\% \times 71.5\%$

[b] $32.9\% = 23.7\% + 9.2\%$

[c] $66.3\% = 23.7\% + 23.4\% + 19.2\%$. Events Graduate and Stage 2-PT Node-Enlisted are dependent as evaluated by P(Enlisted ∩ Not Selected) ≠ P(Enlisted) × P(Not Selected), $23.7\% \neq 32.9\% \times 71.5\% = 23.5\%$

[d] $1.8\% = 7.3\% \times 34.6\% \times 71.5\%$

[e] $2.7\% = 1.8\% + 1.6\% + 1.8\%$

[f] $32.9\% = 1.8\% + 0.9\%$

[g] $35.6\% = 32.9\% + 2.7\%$. Events Graduate and Stage 2-PT Node-Officer are dependent as evaluated by P(Officer ∩ Not Selected) ≠ P(Officer) × P(Not Selected), $1.8\% \neq 2.7\% \times 71.5\% = 1.9\%$

Table 14.9 Stage 1 posterior probabilities

	Status node-stage 1	
Graduate	Enlisted (%)	Officer (%)
Not selected	71.6[a]	70.3
Selected	28.4	29.7
Total	100.0[b]	100.0

Note: This represents the posterior probabilities of the elements of Status Node-Stage 1. The analyst calculated them using the joint and conditional/marginal probabilities reported in Table 14.7 Stage 1-Joint and Marginal Probabilities

[a] $71.6\% = 66.3\%/92.6\%$

[b] $100.0\% = 71.6\% + 28.4\%$

Table 14.10 Stage 2 posterior probabilities

	PT node-enlisted stage 2		
Graduate	Above (%)	Extreme (%)	Average (%)
Not selected	72.0[a]	74.1	68.3
Selected	28.0	25.9	31.7
Total	100.0[b]	100.0	100.0
	PT Node-officer stage 2		
Graduate	Above (%)	Extreme (%)	Average (%)
Not selected	66.7[c]	72.7	72.0
Selected	33.3	27.3	28.0
Total	100.0[d]	100.0	100.0

Note: (1) PT Node-Enlisted Stage 2. This represents the posterior probabilities of the elements of PT Node-Enlisted Stage 2. The analyst calculated them using the joint and conditional/marginal probabilities reported in Table 14.8 Stage 2-Joint and Marginal Probabilities. (2) PT Node-Officer Stage 2. This represents the posterior probabilities of the elements of PT Node-Officer Stage 2. The analyst calculated them using the joint and conditional/marginal probabilities reported in Table 14.8 Stage 2-Joint and Marginal Probabilities

[a]72.0% = 23.7%/32.9%

[b]100.0% = 72.0% + 28.0%

[c]66.7% = 1.8%/2.7%

[d]100.0% = 66.7% + 33.3%

Stage 2. To compute posterior probabilities, the analyst divides the joint probabilities in Table 14.8 Stage 2-Joint and Marginal Probabilities by their respective conditional/marginal probabilities, which totals 100.0%. For example, the analyst computes the posterior probabilities for each element in Stage 2-PT Node-Enlisted and Stage 2-PT Node-Officer by dividing them individually by their respective conditional/marginal probabilities and then reports these in Table 14.10 as posterior probabilities.

Step 10a: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 14.1 using an iterative process where she or he first determines the posterior probabilities from Stage One and then uses these probabilities as the priors for Stage Two.

Step 10b: Draw a tree diagram. The analyst reports posterior probabilities, which they computed by filtering them through the likelihood, joint, and marginal probabilities, which she or he illustrates in Fig. 14.2 using a process called marginalization[2] using the Total Law of Probability.

[2] See Chap. 3 of this book, *Statistical Properties of Bayes' Theorem*, for a discussion of this concept.

Special Forces Assessment and Selection Two Stage Example										
Stage One						Stage Two				
Status Node						PT Node				
Graduate	Status	Likelihood	Joint	Marginal	Posterior	PT	Likelihood	Joint	Marginal	Posterior
						Above	**35.7%**	**23.7%**	**32.9%**	**72.0%**[a]
	Enlisted	**92.7%**	**66.3%**	**92.6%**	**71.6%**[a]	Extreme	35.3%	23.4%	31.6%	74.1%
						Average	29.0%	19.2%	28.1%	68.3%
Not Selected 71.5%										
						Above	34.6%	1.8%		66.7%
	Officer	7.3%	5.2%		70.3%	Extreme	30.8%	1.6%		72.7%
						Average	34.6%	1.8%		72.0%
						Above	**35.0%**	**9.2%**		**28.0%**
	Enlisted	**92.3%**	**26.3%**		**28.4%**	Extreme	31.2%	8.2%		25.9%
						Average	33.8%	8.9%		31.7%
Selected **28.5%**										
						Above	40.9%	0.9%	2.7%	33.3%
	Officer	7.7%	2.2%	7.4%	29.7%	Extreme	27.3%	0.6%	2.2%	27.3%
						Average	31.8%	0.7%	2.5%	28.0%

Fig. 14.1 Tree diagram for the SFAS example. From Tables 14.1, 14.2, 14.3, 14.4, 14.5, 14.6, 14.7, 14.8, 14.9 and 14.10, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities across each Stage of this BBN. [a]If Graduate $= A_i$, Status $= B_j$, and PT $= C_k$, where i = Not Selected, j = Enlisted, and j = Above and invoking BT, $P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{(71.5\%)(92.7\%)}{92.6\%} = 71.6\% = 71.6\% = P(A|B)$. [b]To calculate the posterior probability for stage two, we have $P(A_i|C) = \frac{P(A_i|B)P(C|A_i)}{P(C)} = \frac{(71.6\%)(35.7\%)}{35.5\%} = 72.0\%$ (Note that when the analyst selected Graduate = "Not Selected" and Status = "Enlisted," BT did not require the values for Status = "Officer" ($P(\tilde{B})$) in the calculation of P(A|B) for Stage 1 or the calculation of P(A|BC) for Stage 2, where P(C) = PT = "Above")

Step 11: Run a netica replication[3]. The analyst reports the results of the Netica replication of the prior, conditional, and marginal probabilities of the BBN, which she or he illustrates in Figs. 14.3, 14.4 and 14.5.

14.3 Conclusions

After conducting this experiment, the BNN is loaded with all the available information to date. Now, there is a predictive tool to identify the next occurrence of an event. This tool can evaluate either cause and effect (posterior) relationships or effect and cause (inverse) relationships.

14.3.1 Posterior Probabilities

Given a fully saturated model containing all available information, an analyst can evaluate conditional probability changes going from the cause event, Status, to the

[3] Note that when the analyst selected Graduate = "Not Selected" and Status = "Enlisted," BT did not require the values for Status = "Officer" ($P(\tilde{B})$) in the calculation of P(A|B) for Stage One or the calculation of P(A|BC) for Stage Two, where P(C) = PT = "Above."

Special Forces Assessment and Selection Two Stage Example										
	Stage One					Stage Two				
	Status Node					PT Node				
Graduate	Status	Likelihood	Joint	Marginal	Posterior	PT	Likelihood	Joint	Marginal	Posterior
Not Selected 71.5%	**Enlisted**	**92.7%**	**66.3%**	**92.6%**	**71.6%**	**Above**	**35.7%**	**23.7%**	**32.9%**	**72.0%**[a]
						Extreme	35.3%	23.4%	31.6%	74.1%
						Average	29.0%	19.2%	28.1%	68.3%
	Officer	7.3%	5.2%		70.3%	Above	34.6%	1.8%		66.7%
						Extreme	30.8%	1.6%		72.7%
						Average	34.6%	1.8%		72.0%
Selected **28.5%**	**Enlisted**	**92.3%**	**26.3%**		**28.4%**	**Above**	**35.0%**	**9.2%**		**28.0%**
						Extreme	31.2%	8.2%		25.9%
						Average	33.8%	8.9%		31.7%
	Officer	7.7%	2.2%	7.4%	29.7%	Above	40.9%	0.9%	2.7%	33.3%
						Extreme	27.3%	0.6%	2.2%	27.3%
						Average	31.8%	0.7%	2.5%	28.0%

Fig. 14.2 Tree diagram for the SFAS example. [a]From Tables 14.1, 14.2, 14.3, 14.4, 14.5, 14.6, 14.7, 14.8, 14.9 and 14.10, the analyst can now trace across selected paths in this diagram the respective likelihood, joint, and posterior probabilities of this Three-Event BBN Model. If Graduate = Ai, Status = Bj, and PT = Ck and where we invoke: (1) Not Selected, (2) Enlisted, and (3) Above, we can compute

$$P(A_{Not\,Selected}|B_{Enlisted}\cap C_{Above}) =$$

$$\frac{P(C_{Above}|B_{Enlisted}\cap A_{Not\,Selected})\,P(A_{Not\,Selected}|B_{Enlisted})}{\mathrm{P}(C_{Above}|B_{Enlisted}\cap A_{Not\,Selected})+\mathrm{P}(C_{Above}|B_{Enlisted}\cap A_{Selected})P(A_{Selected}|B_{Enlisted})} =$$

$$(\frac{(35.7\%)(71.6\%)}{(35.7\%)(71.6\%)+(28.4\%)(35.0\%)} = \frac{25.6\%}{25.6\%+9.9\%} = \frac{25.6\%}{35.5\%} = 72.0\%$$

effect event, Graduate. After conducting this experiment and priming the BBN with all available information, following the next officer completing SFAS, an analyst could conclude there is a 28.0% chance that he will and 72.0% that he will not be selected to attend the SFQC. An analyst can obtain similar revised probabilities for enlisted Soldier outcomes. Other events that could contribute to updating the posterior probabilities of this BBN could include if a Soldier is Airborne or Ranger qualified, is married, or if he is a recycle.

14.3.2 Inverse Probabilities

Using inverse probabilities, the analyst can reverse the results above by evaluating conditional probability changes going from the effect event, Status and PT, to the cause event, Graduate. For example in referring to Fig. 14.6, not only can an analyst determine the probability of an Officer with an average PT scores has a 72.0%

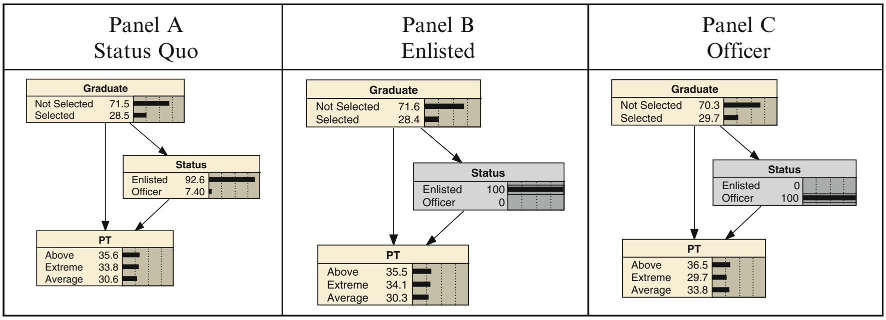

Fig. 14.3 The Netica replication of this SFAS example. *Panel A* represents the subjective or prior probabilities (Graduate Node) and the conditional or marginal joint probabilities (PT and Status Nodes) which are verified in Table 14.1. *Panel B* represents the revised prior or posterior probabilities when an analyst invoked the Event Enlisted; and *Panel C* represents the revised prior or posterior probabilities when an analyst invoked the Event Officer. *Panel B* and *C* probabilities are verified in Stage 1 of Fig. 14.5, P(Not Selected|Enlisted) and P(Not Selected| Officer)

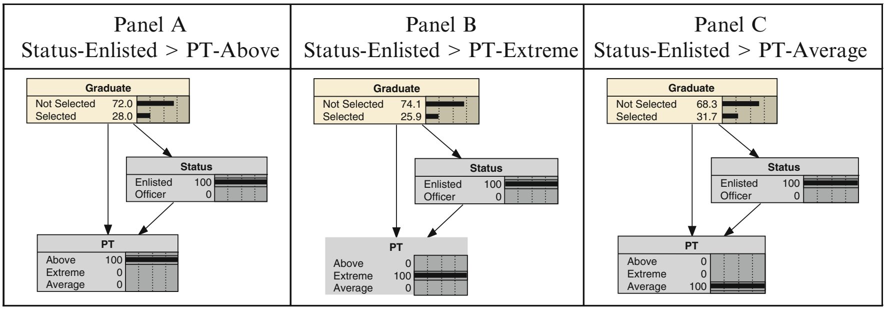

Fig. 14.4 Revised Prior probabilities when the events Enlisted and Above, Extreme, and Average are invoked. The Netica replication of this SFAS example. *Panel A* represents the subjective or prior probabilities (Graduate Node) when an analyst invokes the Events Enlisted and Above P(Not Selected|Enlisted, Officer) = 72.0%. *Panel B* represents the revised prior or posterior probabilities when an analyst invokes the Events Enlisted and Above P(Not Selected|Enlisted, Extreme) = 74.1%. *Panel C* represents the revised prior or posterior probabilities when an analyst invokes the Events Enlisted and Above P(Not Selected|Enlisted, Average) = 68.3%. Table 14.10 verifies these probabilities

chance of not being selected, as seen above, P(Not Selected|Officer, Average), but she or he can also determine the percentage effect of Soldiers that have above average PT scores and who are officers who will not be selected, P(Above, Officer| Not Selected), which is 35.7% (Above) and 7.27% (Officer), respectively. Invoking Selected will similarly adjust these conditional probabilities.[4]

[4] Due to very low selection rates and the smaller proportion of officers attending SFAS, this percentage is also low.

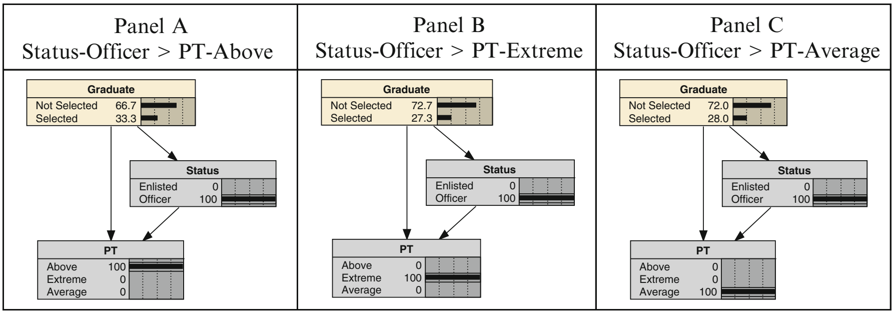

Fig. 14.5 Revised Prior probabilities when the events Officer and Above, Extreme, and Average are invoked. The Netica replication of this SFAS example. *Panel A* represents the subjective or prior probabilities (Graduate Node) when an analyst invokes the Events Officer and Above P(Not Selected|Enlisted, Officer) = 66.7%. *Panel B* represents the revised prior or posterior probabilities when an analyst invokes the Events Enlisted and Above P(Not Selected|Officer, Extreme) = 72.7%. *Panel C* represents the revised prior or posterior probabilities when an analyst invokes the Events Officer and Above P(Not Selected|Officer, Average) = 72.0%. These probabilities are verified in Table 14.10

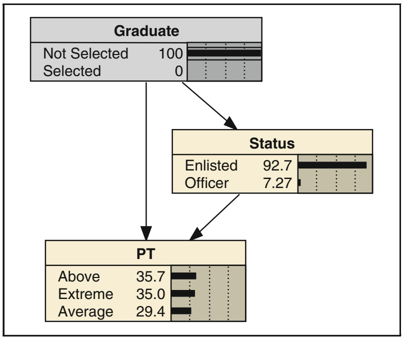

Fig. 14.6 Represents the effects on the conditional probabilities using inverse probability when an analyst inverts the cause and effect relationship. For example in referring to Fig. 14.6, not only can an analyst determine the probability of an Officer with an average PT scores has a 72.0% chance of not being selected, as seen above, P(Not Selected|Officer, Average), but she or he can also determine the percentage effect of Soldiers that have above average PT scores and who are officers who will not be selected, P(Above, Officer|Not Selected), which is 35.7% (Above) and 7.27% (Officer), respectively. Invoking Selected will similarly adjust these conditional probabilities

References

Special Forces (2012). http://www.sorbrecruiting.com/. Accessed 2 July 90 2012.

Index

J. Grover, *Strategic Economic Decision-Making: Using Bayesian Belief Networks to Solve Complex Problems*, SpringerBriefs in Statistics 9,
DOI 10.1007/978-1-4614-6040-4, © Springer Science+Business Media New York 2013