Chapter 16: Bayesian Inference and Human Inference

In this chapter you will learn about Bayes theorem, a rational procedure for revising beliefs. You will learn how to set up a Web-based experiment that tests if the theorem is descriptive of human inference, and you will also learn how to analyze the data from such an experiment.

If you toss a coin, what is the probability that it will be heads? What is the probability that a major league baseball player will hit more than 68 home runs in the next baseball season?

The first question dealing with tossing a coin refers to an experiment that can be repeated. One assumes that the coin has no memory, and that it will behave in the same way from day to day and year to year. One can develop and test a physical theory that the two sides of the coin are equally likely to fall facing up, independent of history and the actions of other coins. Evidence from the past can be used to infer the probability of what will happen in the future. The second question (home runs) refers to a unique event that will either happen or will not happen, and there is no way to calculate a proportion that is clearly valid. One might look at what other ball players have done and can ask if those who are likely to accomplish the feat seem healthy, but players change, the conditions (e.g., pitchers) change, and it is never really the same experiment. This type of situation is sometimes referred to as one of *uncertainty*, and the term *subjective probability* is sometimes used to refer to the psychological strength of belief that the event will happen.

Nevertheless, people are willing to use the same term, probability, to express both types of ideas. People are willing to gamble on both types of predictions—on repeatable, mechanical games of chance (like dice, cards, and roulette) and on unique events (like horse races and sporting contests). In fact, people are even willing to use the same language *after* something has happened (a murder, for example) to discuss the "probability" that a particular event occurred (e.g., this defendant committed the crime).

The Rev. Thomas Bayes (1702-1761) derived a theorem for inference from the mathematics of probability. Philosophers recognized that this theorem could be interpreted as a calculus for rational thought. Psychologists have investigated if Bayes theorem also describes how people form and revise their beliefs (Birnbaum, 1983; Birnbaum & Mellers, 1983; Edwards, 1968; Gigerenzer & Hoffrage, 1995; Koehler, 1996; Novemsky & Kronzon, 1999; Shanteau, 1975a; Troutman & Shanteau, 1977; Wallsten, 1972).

A. Bayes Theorem

To illustrate how Bayes theorem works, it helps to work through an example. Suppose there is a disease that infects one person in 1000, completely at random. Suppose also that there is a blood test for the disease that yields a “positive” test result in 99.5% of cases of the disease and gives a false “positive” in only 0.5% of people who do not have the disease. Suppose a person in this study tests “positive,” what is the probability that he or she has the disease? Think for a moment and see if you can intuitively arrive at a judgment of the likelihood that a person who tests “positive” is actually sick with the disease in this case. The solution, according to Bayes theorem, may surprise you.

Before you read on, you should load the experiment, *Bayes.htm*, from the CD, and participate as a judge in the experiment on the *Cab Problem*. This experiment was constructed with the help of factorWiz. Later in this chapter, you will learn to analyze data from this experiment and compare them with the theory of Bayes.

Suppose there are two hypotheses, H and not-H (denoted H'). In the example above, they are hypothesis that the person is sick with the disease (H) and the complementary hypothesis (H') that the person does not have the disease. Let D refer to the datum that is relevant to the hypotheses.

In the example, D would refer to a “positive” test result, and D' would be a “negative” result from the blood test.

The problem stated that 1 in 1000 have the disease, so P(H) = .001; i.e., the prior probability (before we test the blood) that a person has the disease is .001, and the probability that the person is not sick with the disease, is .999 = P(H’) = 1 – P(H).

The problem also gave information about the diagnostic value of the test. The conditional probability that a person will test “positive” given that person has the disease is written as P(“positive”| H) = .995, and the conditional probability that a person will test “positive” given he or she is not sick is P(“positive”| H’) = .005. These two probabilities are also called *the HIT RATE* and *FALSE ALARM RATE* in signal detection, and they are also known as *power* and *significance* (α) in statistics. We need to calculate P(H| D), the probability that a person is sick, given the test was “positive.” This calculation is known as an inference.

The conditional probability of A given B is not the same as the conditional probability of B given A. For example, the probability that someone is male given he or she is a member of the U.S. senate is quite high because there are few women in the senate. However, the probability that a person is a member of the U.S. senate given that person is male is quite low, since there are so few senators and so many males. Conditional probability is defined as follows:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (16.1)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad (16.2)$$

So, if A is the set of U.S. Senators (of which there are 100), and B is the set of males (of which there are billions), we see that the probability of A given B will be quite small, but the probability of B given A can be quite high.

The situation is as follows: we know P(H), P(D|H) and P(D|H'), and we want to calculate P(H|D). From the definition of conditional probability:

$$P(H \mid D) = \frac{P(H \cap D)}{P(D)} \tag{16.3}$$

Also, from the definition of conditional probability, $P(H \cap D) = P(D \mid H)P(H)$. In addition, D can happen in two mutually exclusive ways, either with H or without it, so $P(D) = P(D \cap H) + P(D \cap H')$. Each of these conjunctions can be written in terms of conditionals, so by substitution the formula is as follows:

$$P(H \mid D) = \frac{P(D \mid H)P(H)}{P(D \mid H)P(H) + P(D \mid H')P(H')} \tag{16.4}$$

Equation 16.4 is known as Bayes Theorem. Substituting the values given for the blood test problem yields the following result:

$$P(sick \mid \text{"}positive\text{"}) = \frac{(.995)(.001)}{(.995)(.001) + (.005)(.999)} = .166.$$

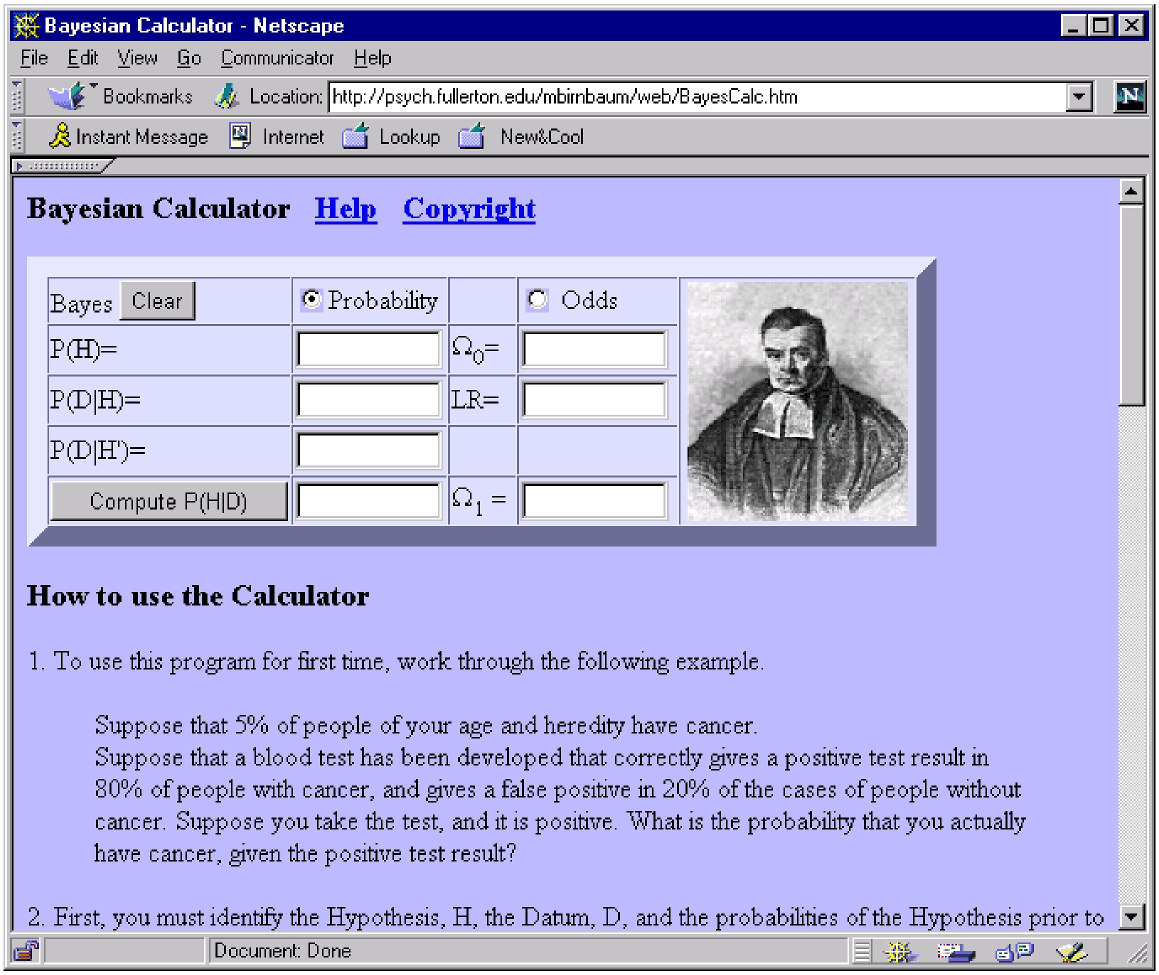
Does this result surprise you? Think of it this way: Among 1000 people, there is only one sick person. If all 1000 were given the test, the test would probably give a "positive" test result to that one person, and it would also give a "positive" test result to about five others (of the 999 healthy people 0.5% should test positive). Thus, of the six people who test "positive," only one is really sick, so the probability of being sick, given a "positive" test result, is only about one in six. Another way to look at the .166 is that it is 166 times bigger than the probability of being sick given no information about the person (.001), so there has indeed been considerable revision of opinion given the positive test.

B. Bayesian Calculator

The calculations of Equation 16.4 can be facilitated by the Bayesian calculator included on your CD. You can load it from the index of examples for this chapter. To use the calculator, first you need to define H and D. Then identify and type in P(H), P(D|H), P(D|H') into the first column. (Leave the second column and the last row value blank). Press the calculate button, and the answer, P(H|D), will appear in the last row. Test the calculator with the disease problem given above. Instructions are also given in the Web page for the calculator.

If you view the page source of the calculator, you will see that the first half uses HTML forms used to define the table, and the second half is a JavaScript program that does the calculations. JavaScript will be explained Chapters 17-19. The appearance of the calculator is shown in Figure 16.1. Insert Figure 16.1 about here.

Figure 16.1. The Bayesian Calculator. This calculator is written in JavaScript, which will be explained in Chapters 17-18. Rev. Thomas Bayes is pictured on the calculator.



C. The Cab Problem

The experiment on the Cab problem is based on a problem used by Tversky and Kahneman (1982; see also Kahneman & Tversky, 1973). Instructions for the Cab problem on your CD read (in part) as follows:

"A cab was involved in a hit-and-run accident at night. There are two cab companies in the city, the Blue and Green. Your task is to judge (or estimate) the probability that the cab in the accident was a Blue cab. You will be given information about the percentage of accidents at night that were caused by Blue cabs, and the testimony of a witness who saw the accident.

"The percentage of night-time accidents involving Blue cabs is based on the previous 2 years in the city. In different cities, this percentage was either 15%, 30%, 70%, or 85%. The rest of night-time accidents involved Green cabs.

"Witnesses were tested for their ability to identify colors at night. The witnesses were presented 100 Blue and 100 Green cabs to identify at night.

"The MEDIUM witness correctly identified 60% of the cabs of each color, calling Green cabs "Blue" 40% of the time and calling Blue cabs "Green" 40% of the time.

"The HIGH witness correctly identified 80% of each color, calling Blue cabs "Green" or Green cabs "Blue" on 20% of the tests. "

When you have completed the task, use the calculator to find the Bayesian calculations assuming that the prior probabilities are P(H) = .15, .30, .70, or .85, where H = the hypothesis that the cab was Blue. Suppose the Medium credibility witness has P("B"|H) = .6 and P("G"|H') = .4; and suppose the High credibility witness has P("B"|H) = .8 and P("B"|H') = .2. For example, if 15% of accidents involve Blue cabs and the High credibility witness says the cab was "Green," the Bayesian probability that it was a Blue cab is only .04. The datum that the witness said the cab was "Green" tends to exonerate the Blue cab. To find that value, realize that the probability that the witness would say "Green" given the cab was Blue is only .2 and the probability that the witness would say "Green" given it was Green is .8. In the calculator, enter .15 for P(H), .2 for P(D|H) and .8 for P(D|H').

Suppose there are 15% Blue cabs but the High credibility witness said "Blue." In this case, the Posterior Probability (Probability it was Blue given witness said "Blue") is .413.

There are additional complications to the Bayesian solution to the cab problem. One involves the extrapolation of the witness's performance during the test to the conditions on the night of the accident (Birnbaum, 1983). Another involves the inference from the percentage of accidents caused by Blue cabs in some period to the prior probability that the cab was Blue. In a later section, a more general, subjective Bayesian model will be presented that allows such complications.

D. Extract the Data for the Cab Problem and Find Means

The first step is to filter the data. In Excel, open the data file, *clean.xls*, from the CD. Type variable names in the first row, if you have not done so already. Click in Cell A1, and select *Filter:Autofilter* from the **Data** menu. Then click the drop down selection arrow in the A1 and select *bayes*. Copy the data to a new file, and save the file as an Excel Workbook with the name *bayes.xls*. (A completed file of *bayes.xls* is included on the CD to allow you to check your work.)

The variable names are included at the end of the *bayes* data, so cut them and copy them in the first row. To make the variable names, take the experiment and type in the stimuli rather than answers. Then, in NotePad, replace the date, time, and sex with the names of these variables.

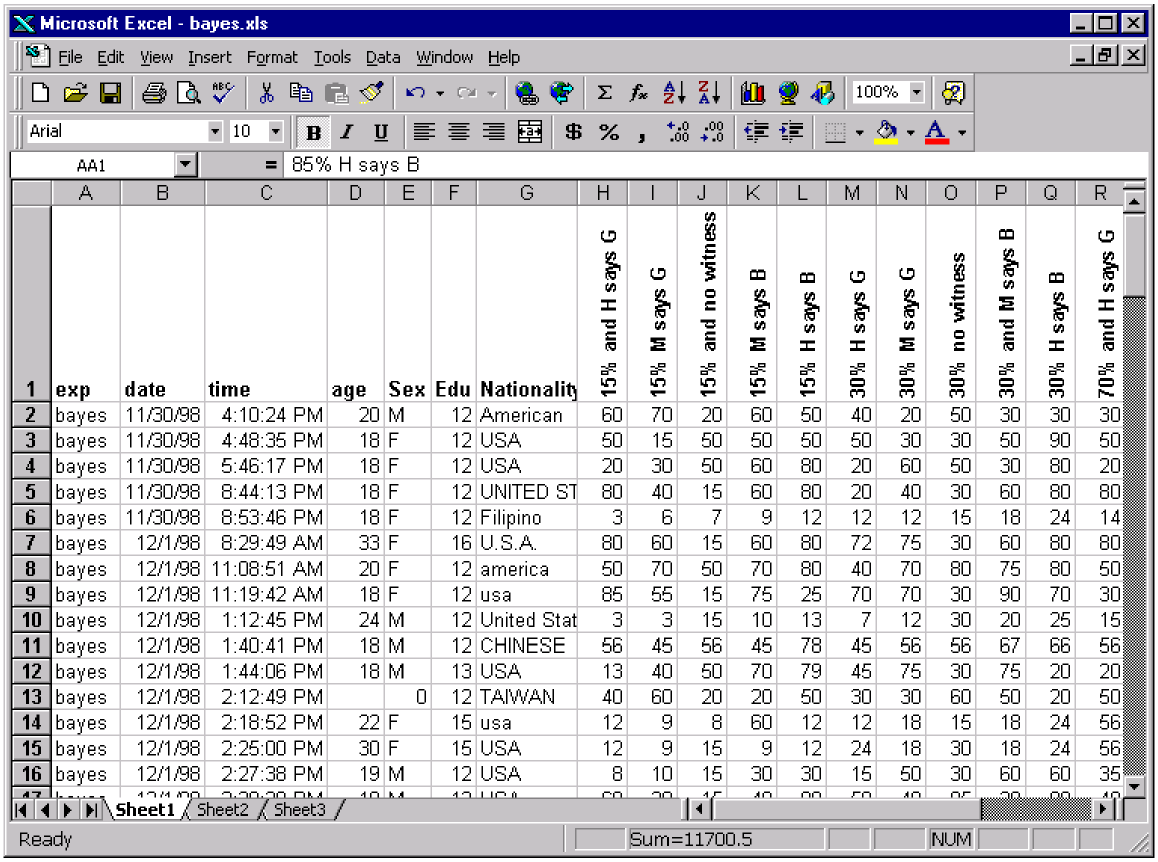
The second step is to remove blank lines and multiple submissions from the data. Use conditional formatting to check for responses that are out of bounds (e.g., negative numbers or values above 100. The techniques for conditional formatting are described in Chapter 12. The resulting data file is shown in Figure 16.2. Insert Figure 16.2 about here.

To find means, click in the row after the last row of data in Column H, and click *Insert: function: Average* to find the mean judgment of probability for the case of 15% Blue cabs have accidents at night and a High credibility witness says that the cab was "Green." Then use *AutoFill* to find the other column means, using the technique described in previous chapters.

The third step is to use *Copy* and *Paste Special* to copy the means to a new worksheet in the same workbook. Be sure to keep the means linked so that any change in the data will be reflected in the rest of the workbook. Then cut and paste to arrange the means in matrix form. This process was also described in previous chapters. Next, add labels for the rows and columns. The matrix of means is shown in Figure 16.3.

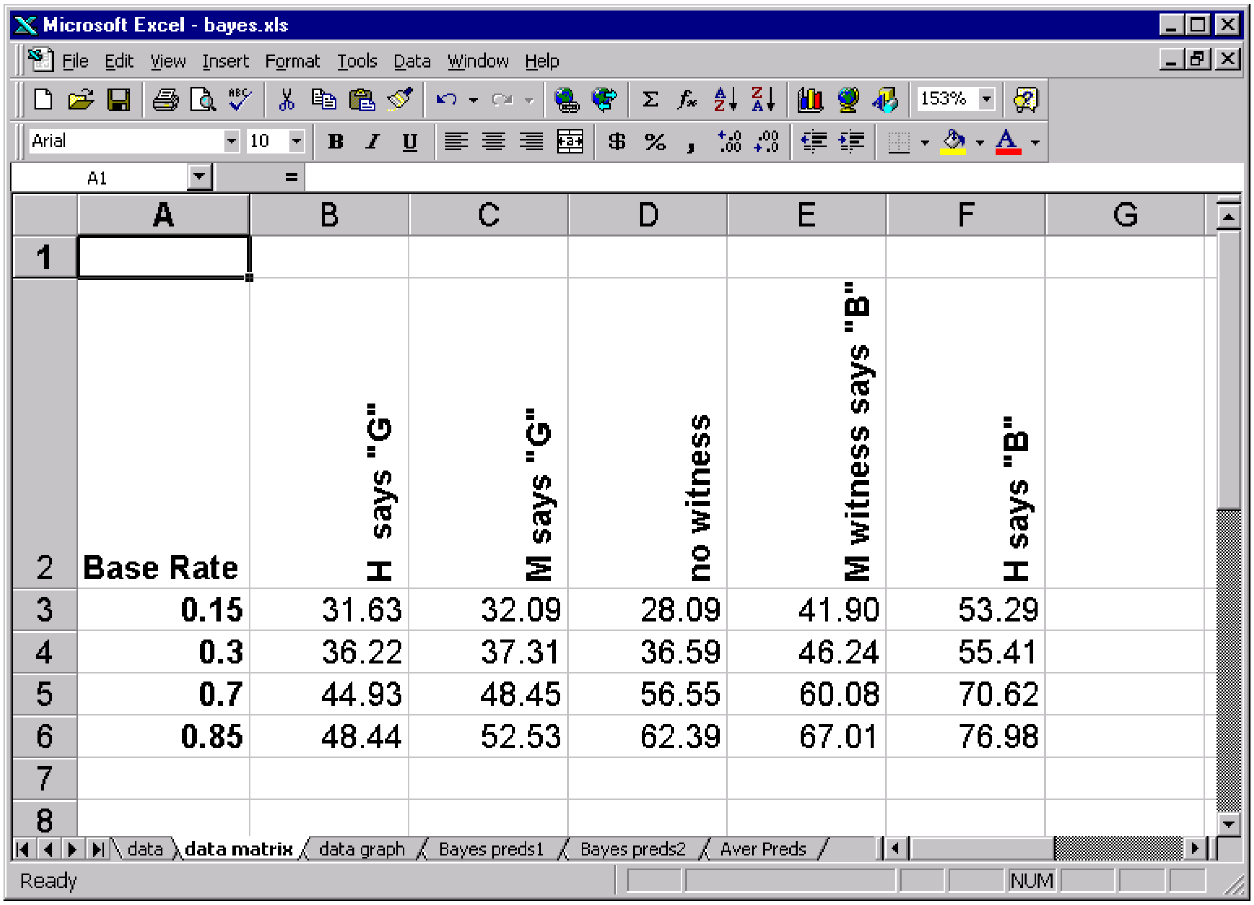
Insert Figure 16.3 about here.

Figure 16.2. Appearance of data in file, *bayes.xls*.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	exp	date	time	age	Sex	Edu	Nationality	15% and H says G	15% M says G	15% and no witness	15% M says B	15% H says B	30% H says G	30% M says G	30% no witness	30% and M says B	30% H says B	70% and H says G
2	bayes	11/30/98	4:10:24 PM	20	M	12	American	60	70	20	60	50	40	20	50	30	30	30
3	bayes	11/30/98	4:48:35 PM	18	F	12	USA	50	15	50	50	50	50	30	30	50	90	50
4	bayes	11/30/98	5:46:17 PM	18	F	12	USA	20	30	50	60	80	20	60	50	30	80	20
5	bayes	11/30/98	8:44:13 PM	18	F	12	UNITED ST	80	40	15	60	80	20	40	30	60	80	80
6	bayes	11/30/98	8:53:46 PM	18	F	12	Filipino	3	6	7	9	12	12	12	15	18	24	14
7	bayes	12/1/98	8:29:49 AM	33	F	16	U.S.A.	80	60	15	60	80	72	75	30	60	80	80
8	bayes	12/1/98	11:08:51 AM	20	F	12	america	50	70	50	70	80	40	70	80	75	80	50
9	bayes	12/1/98	11:19:42 AM	18	F	12	usa	85	55	15	75	25	70	70	30	90	70	30
10	bayes	12/1/98	1:12:45 PM	24	M	12	United Stat	3	3	15	10	13	7	12	30	20	25	15
11	bayes	12/1/98	1:40:41 PM	18	M	12	CHINESE	56	45	56	45	78	45	56	56	67	66	56
12	bayes	12/1/98	1:44:06 PM	18	M	13	USA	13	40	50	70	79	45	75	30	75	20	20
13	bayes	12/1/98	2:12:49 PM		0	12	TAIWAN	40	60	20	20	50	30	30	60	50	20	50
14	bayes	12/1/98	2:18:52 PM	22	F	15	usa	12	9	8	60	12	12	18	15	18	24	56
15	bayes	12/1/98	2:25:00 PM	30	F	15	USA	12	9	15	9	12	24	18	30	18	24	56
16	bayes	12/1/98	2:27:38 PM	19	M	12	USA	8	10	15	30	30	15	50	30	60	60	35

Figure 16.3. Mean judgments of the probability (expressed as a percentage) that the cab was Blue, given the Base rate (of blue cabs involved in accidents at night) and the testimony of a witness (who said the cab was either “Blue” or “Green.”) H = high credibility witness; M = medium credibility witness.



Base Rate	H says "G"	M says "G"	no witness	M witness says "B"	H says "B"
0.15	31.63	32.09	28.09	41.90	53.29
0.3	36.22	37.31	36.59	46.24	55.41
0.7	44.93	48.45	56.55	60.08	70.62
0.85	48.44	52.53	62.39	67.01	76.98

E. Graph the Data

To graph the data, select Cells A2:F6, then select *Chart* from the **Insert** menu. From the Chart wizard, choose *XY scatter*, with data points connected by lines. In the second step of the chart wizard, specify that the data are in columns (that will plot the data as a function of the prior probability, or base rate, of Blue cabs). After suitable labels have been added for the axes, and adjustments have been made in the fonts, marker styles, and other features of the graph, the graph appears as in Figure 16.4. Insert Figure 16.4 about here.

Clearly, people use both the base rate and the witness information. The spread between curves represents the effect of the witness testimony, and the slopes of the curves represent the effect of the base rate. Research on the "base rate fallacy" contended that people do not attend to base rates (Kahneman & Tversky, 1973), based on a nonsignificant result in a between-subjects design. If people did not attend to the base rate, the curves in Figure 16.4 would haveslopes of zero; i.e., they would be flat. However, Figure 16.4 shows that people do attend to base rate, even in the cab problem, when base rates are manipulated within-subjects. As noted in Chapter 9, it can be tricky to compare data between groups of subjects who judge different stimuli. Birnbaum and Mellers (1983) studied the same Bayesian problem within- and between-subjects and found quite different results. The data in Figure 16.4 are within-subjects, and they show that people utilize the base rate.

Note that the curve for *no witness* is steeper than the other curves, and even crosses over two of the other curves. As you will learn in a later section, this crossover is not consistent with additive models (Novemsky & Kronzon, 1999), including a subjective version of Bayes theorem (Wallsten, 1972). Instead, the crossover is indicative of an averaging model. To understand how the data differ from Bayesian theory, you will graph predictions for Bayes theorem.

F. Compute Bayes Theorem

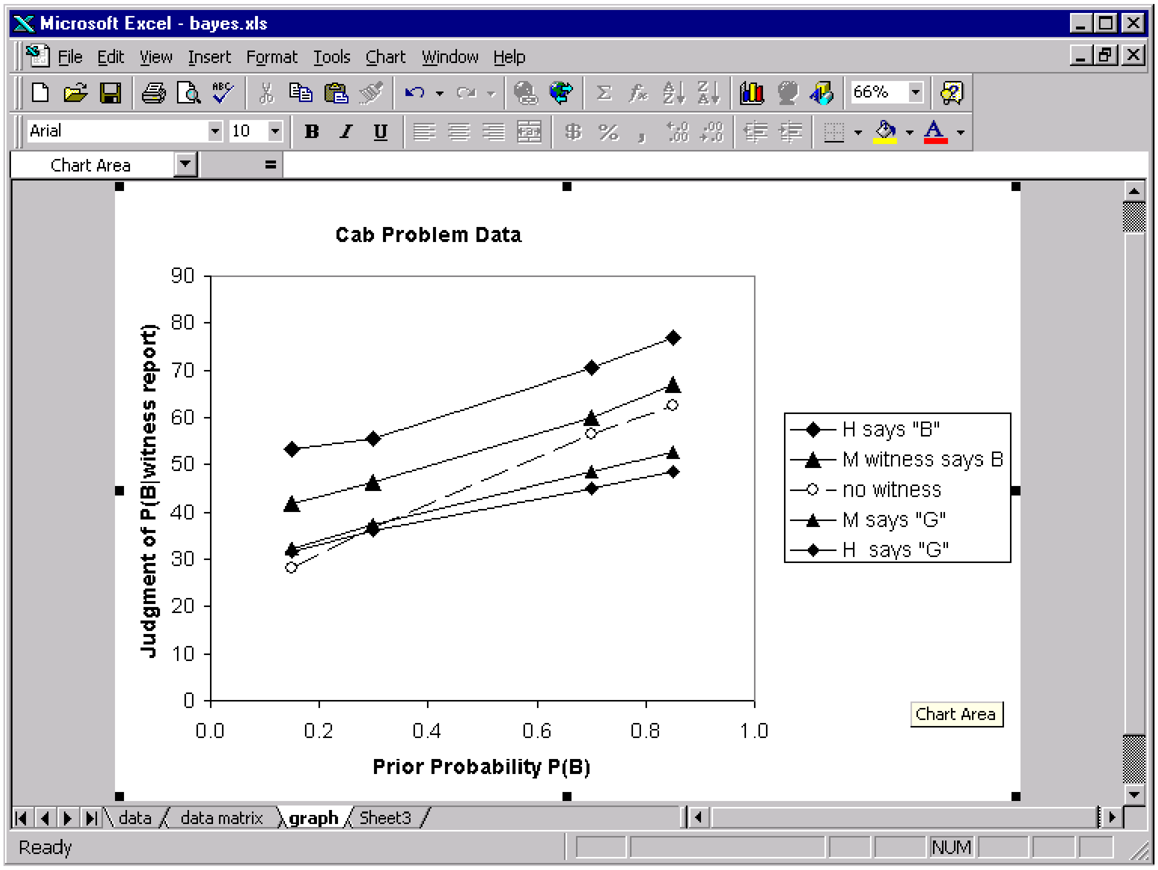
To construct predictions for Bayes Theorem, type in the base rates, .15, .3, .7, and .85 in Cells H3:H6 of the worksheet with the matrix of means. Select those cells, and type in the name *BR* (for Base rate) in the name box, and hit enter. Then type in Cells I2:M2 the values of P(D|H) = .2, .4, .5, .6, and .8 for the H witness who says “G”, for M witness who says “G”, no witness, M witness who says “B”, and H witness who says “B.” Select all cells in I2:M2 and type *PDGH* in the *Name Box* (Probability of the Datum Given the Hypothesis), then hit enter. Next, in Cells I1:M1 type the values of P(D|H’) = .8, .6, .5, .4, and .2; name them *PDGNH* (for Probability of the Datum Given Not the Hypothesis). In Cell I3, type the following expression (Bayes Theorem):

*=(PDGH*BR)/(PDGH*BR+PDGNH*(1-BR))*

Next hit enter, and the value .042 will appear in Cell I3. Next, click the mouse pointer in the lower right corner of cell I3 until you see the AutoFill “+”. Drag down to fill in the first column, and while the column is selected, drag to the right to AutoFill the matrix. If you did everything right, the matrix will appear as in Figure 16.5. If you have a problem, check to see the names you gave the variables, and check the expression typed in Cell I3 very carefully. You can undo your most recent steps with the *Undo* function (**Edit** menu); so when you catch an error, keep selecting *Undo* until you have removed the error. Then retype the correct expressions.

Insert Figure 16.5 about here.

Figure 16.4. Figure of the mean judgments for the Cab Problem. Judgments of the P("Blue"|witness testimony) are plotted as a function of the prior probability of Blue, with a separate curve for each level of the witness testimony.

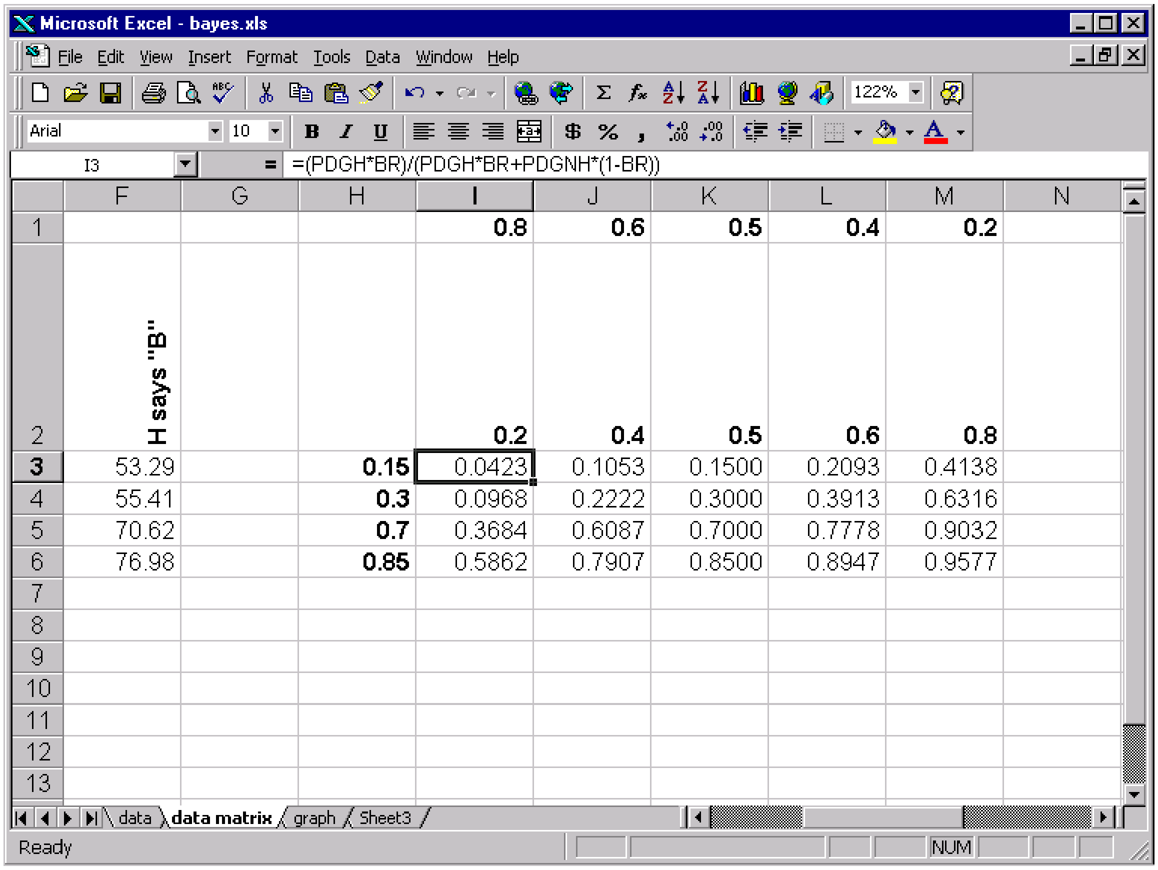


G. Graph Predictions of Bayes Theorem

The next step is to plot the predictions in the same manner as was done for the data. The predictions are shown in Figure 16.6. Insert Figure 16.6 about here. The predictions, shown in Figure 16.6 can be compared to the data shown in Figure 16.4. This comparison shows that judgments are “conservative” relative to Bayes Theorem. If 85% of night time accidents are caused by Blue cabs and the High witness says the cab was “Blue”, Bayes Theorem predicts a value of .96, but the mean judgment is only 78% (.78). Only 7 of 153 people gave responses greater than .95 for this case. Probabilities predicted to be close to zero (when P(B) = .15 and the High witness said “Green”) are also regressed toward .5. The Bayesian prediction is .04 and the mean response is 31.6% (.32). This failure to give judgments as extreme as those calculated by Bayes Theorem was described as “conservatism” in early literature, reviewed by Edwards (1968).

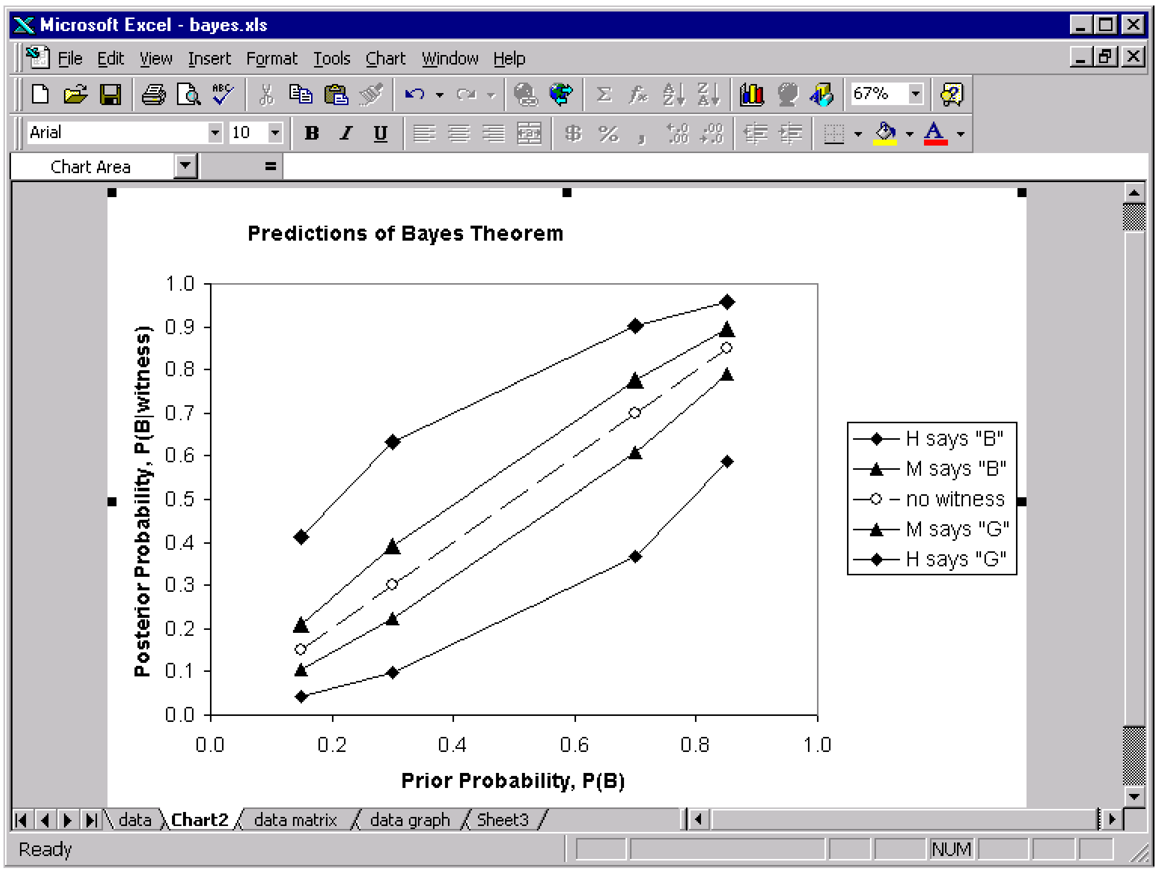
Another aspect of the data can be seen in the following comparions. When the base rate is .85 and the High witness says “Green,” the prediction is .58; when the base rate is .15 and the H witness says “Blue,” the prediction is .41; Bayes theorem says the .15 base rate outweighs a High credibility (.8) source. However, the mean judgments in these two cases have the opposite relation (48% to 53%). This pattern might be indicative of underweighting of base rate information, a less extreme form of the type of base rate neglect argued for by Tversky and Kahneman (1982).

Figure 16.5. Bayesian calculations. Each entry is Excel's calculation according to Bayes Theorem (the expression typed in Cell I3, shown in the formula line). Check these calculations with the Bayesian Calculator.



	F	G	H	I	J	K	L	M	N
1				0.8	0.6	0.5	0.4	0.2	
2	H says "B"			0.2	0.4	0.5	0.6	0.8	
3	53.29		0.15	0.0423	0.1053	0.1500	0.2093	0.4138	
4	55.41		0.3	0.0968	0.2222	0.3000	0.3913	0.6316	
5	70.62		0.7	0.3684	0.6087	0.7000	0.7778	0.9032	
6	76.98		0.85	0.5862	0.7907	0.8500	0.8947	0.9577	
7									
8									
9									
10									
11									
12									
13									

Figure 16.6. Predictions of Bayes Theorem, assuming that P(B) are the values for the previous two years and that P(D|H) and P(D|H') are the same as the values in the previous two years. The comparison of predictions like these and data as shown in Figure 16.4 led to the conclusion that people are "conservative" relative to Bayes Theorem. In other words, people deviate less from .5 than they are predicted to do by use of these values.

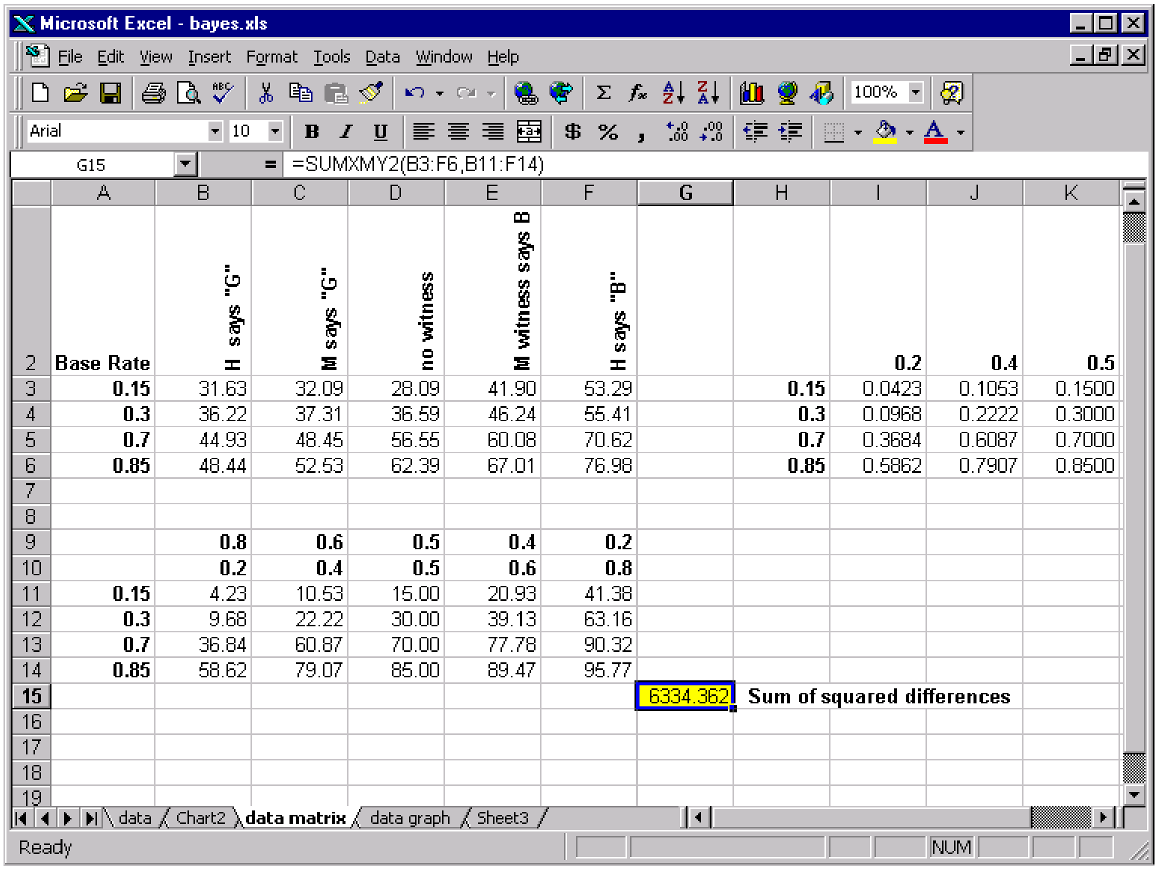


H. Subjective Bayesian Theory

Perhaps people use Bayesian reasoning, but they do not assign their subjective priors to the values given in the problem for Blue cabs involved in accidents at night. In addition, they may use different subjective values for the hit rate and false alarm rates of the witnesses from those specified in the problem. Indeed, the test used an equal number of Blue and Green cabs, but on the night of the accident, the witness may have been aware of the base rate and taken that into account. Wallsten (1972) suggested such a subjective version of the Bayesian model. A monotonic transformation converts the subjective Bayesian model into an additive model (Birnbaum & Mellers, 1983; Novemsky & Kronzon, 1999).

We can represent a Bayesian process with subjective values by using the equation of Bayes Theorem and solving for the values of the evidence. To do this, multiply the predictions by 100 (to express them as percentages, comparable to the data), and take the difference between data and predictions. Next, use the solver to minimize the sum of squared differences between data and predictions. In Cells A9:F14 construct the same matrix as in the previous analysis, except multiply each entry by 100. Next, click in Cell G15 (any convenient cell would do), and from the **Insert** menu, choose *Function*. Select SUMXMY2, the function that computes the sum of squared differences between two arrays. Select the data for the first array (B3:F6). Then specify the predictions for the second array (B11:F14). The sum of squared differences is 6334.4. This step is shown in Figure 16.7. Insert Figure 16.7 about here.

Figure 16.7. Predictions and Data. The selected cell shows the sum of squared differences prior to use of the Solver.



Microsoft Excel - bayes.xls

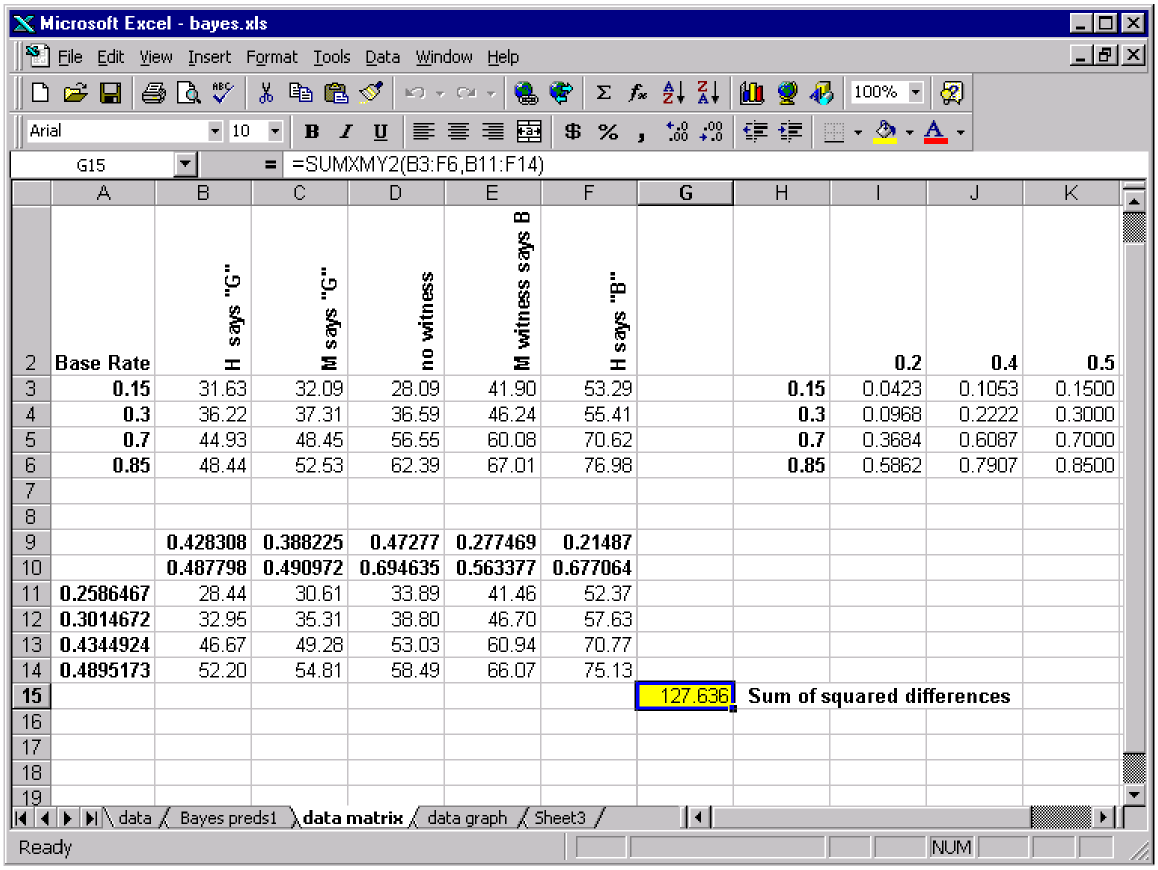
G15 = =SUMXMY2(B3:F6,B11:F14)

	A	B	C	D	E	F	G	H	I	J	K
2	**Base Rate**	**H says "G"**	**M says "G"**	**no witness**	**M witness says B**	**H says "B"**			**0.2**	**0.4**	**0.5**
3	**0.15**	31.63	32.09	28.09	41.90	53.29		**0.15**	0.0423	0.1053	0.1500
4	**0.3**	36.22	37.31	36.59	46.24	55.41		**0.3**	0.0968	0.2222	0.3000
5	**0.7**	44.93	48.45	56.55	60.08	70.62		**0.7**	0.3684	0.6087	0.7000
6	**0.85**	48.44	52.53	62.39	67.01	76.98		**0.85**	0.5862	0.7907	0.8500
7											
8											
9		**0.8**	**0.6**	**0.5**	**0.4**	**0.2**					
10		**0.2**	**0.4**	**0.5**	**0.6**	**0.8**					
11	**0.15**	4.23	10.53	15.00	20.93	41.38					
12	**0.3**	9.68	22.22	30.00	39.13	63.16					
13	**0.7**	36.84	60.87	70.00	77.78	90.32					
14	**0.85**	58.62	79.07	85.00	89.47	95.77					
15							6334.362	**Sum of squared differences**			

data / Chart2 / data matrix / data graph / Sheet3

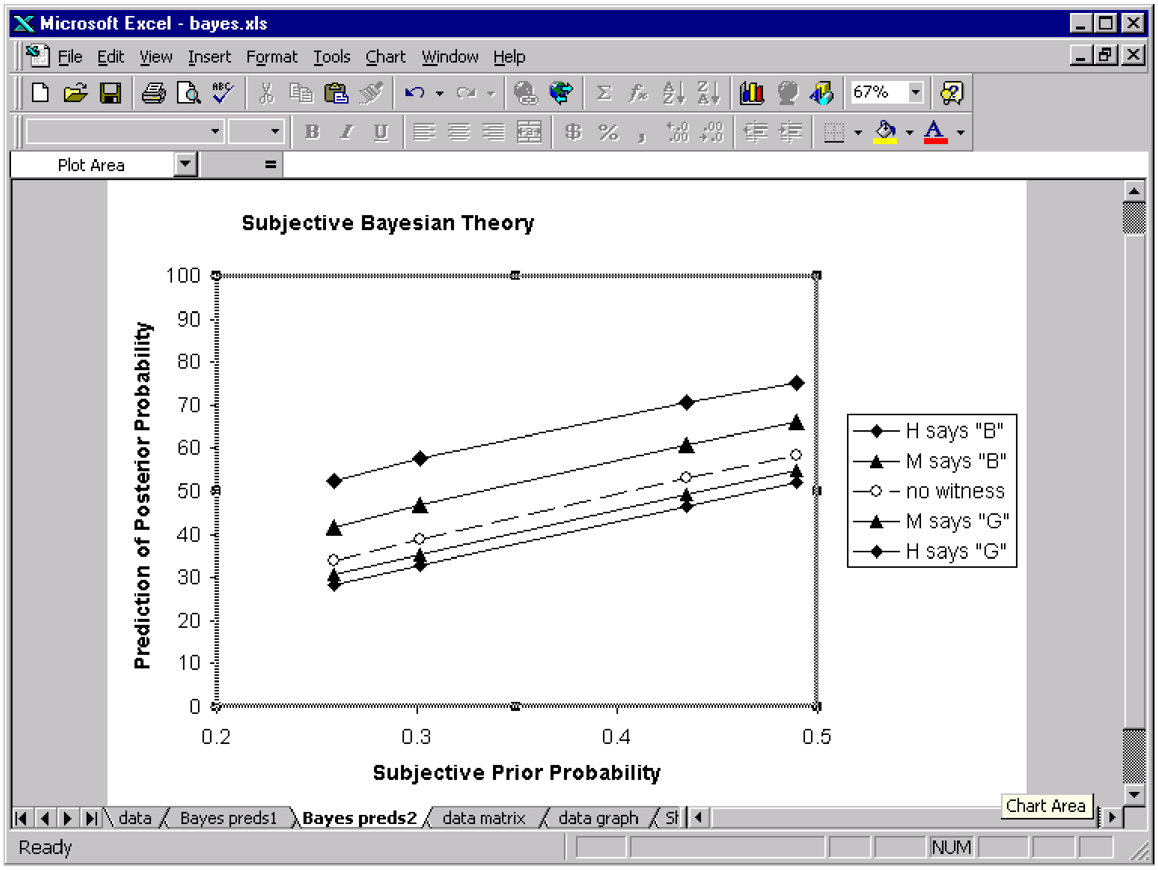
Ready NUM

Figure 16.8. Solution to the Subjective Bayesian Theory, after the *Solver* has done its work. Note that the subjective values of the priors are .26 to .49. Note also the contrast between the subjective values of the hit and false alarm rates and the objective values in Figure 16.7.



	A	B	C	D	E	F	G	H	I	J	K
2	Base Rate	H says "G"	M says "G"	no witness	M witness says B	H says "B"			0.2	0.4	0.5
3	0.15	31.63	32.09	28.09	41.90	53.29		0.15	0.0423	0.1053	0.1500
4	0.3	36.22	37.31	36.59	46.24	55.41		0.3	0.0968	0.2222	0.3000
5	0.7	44.93	48.45	56.55	60.08	70.62		0.7	0.3684	0.6087	0.7000
6	0.85	48.44	52.53	62.39	67.01	76.98		0.85	0.5862	0.7907	0.8500
7											
8											
9		0.428308	0.388225	0.47277	0.277469	0.21487					
10		0.487798	0.490972	0.694635	0.563377	0.677064					
11	0.2586467	28.44	30.61	33.89	41.46	52.37					
12	0.3014672	32.95	35.31	38.80	46.70	57.63					
13	0.4344924	46.67	49.28	53.03	60.94	70.77					
14	0.4895173	52.20	54.81	58.49	66.07	75.13					
15							127.636	Sum of squared differences			

Figure 16.9. Graph of the Subjective Bayesian Theory. Although these predictions are closer to the data than the previous predictions, they do not explain the curve for no witness.



The next step is to select *Solver* from the Tools menu. Select Cell G15 as the target cell; choose to *Minimize* (click the radio button next to *Min*); click in the box for *By changing cells* and select the prior probabilities, hit and false alarm rates (i.e., Cells A11:A14 and B9:F10). After the Solver has worked, the cells change and the sum of squared differences changes, as shown in Figure 16.8. The new predictions, shown in Figure 16.9, are closer to the data, but still do not predict the crossover of the dashed curve, representing judgments when there was no witness.

Insert Figure 16.8-16.9 about here.

I. Averaging model of Source Credibility

Another theory is that people aggregate probabilistic information by an averaging process, rather than by Bayes theorem in either the objective, subjective, or additive form (Anderson, 1974; Shanteau, 1975a; Birnbaum & Mellers, 1983). An averaging model can be written as follows:

$$R = \frac{\sum_{i=0}^{n} w_i s_i}{\sum_{i=0}^{n} w_i} \qquad (16.5)$$

where R is the subjective impression of probability, w_i and s_i are the weight and scale value of source of information i, and w_0 and s_0 are the weight and scale value of the initial impression. The base rate information should have a single weight and scale values that depend on the percentage of Blue cabs in accidents at night. The scale value and weight of the sources depend on the hit rates and false alarm rates of the sources (Birnbaum & Mellers, 1983). The greater the difference between hit rate and false alarm rate, the greater the weight. The scale value depends on whether the source said the cab was “Blue” or “Green” and the tendency of the witness (bias) to say cabs are “Blue” or “Green” (the sum of hit rate and false alarm rate, which was constant in this study). When no witness is presented, the weight of the witness is zero. The averaging model predicts that the effect of the base rate is greater when there is no witness because the weight of the witness drops out of the denominator, increasing the relative weight of the base rate.

You can also use Excel to fit this theory. Figure 16.10 shows the spreadsheet set up with weights and scale values. Note that the equations in the Cells are different, depending on whether the source said “Green”, “no witness”, or “Blue.” The variable names assigned in Excel are w_0 =W_0; s_0 = S_0, the weight of a source is W_S. The scale value of “Green” is S_G, the scale value of “Blue” is S_B, scale value of the base rate is S_BR, and the weight of the base rate is W_B. The expression when the source says “Green” is

=100*(W_0*S_0+W_B*S_BR+W_S*S_G)/(W_0+W_B+W_S)

when the source says “Blue”, the expression is the same, except the scale value of S_B is substituted, as follows:

=100*(W_0*S_0+W_B*S_BR+W_S*S_B)/(W_0+W_B+W_S)

When there is no witness, the weight of W_S is set to zero, so the expression simplifies as follows:

=100*(W_0*S_0+W_B*S_BR)/(W_0+W_B)

Type in these expressions, and use *AutoFill* again to compute the predictions in each respective sub-portion of the matrix. Next, compute the sum of squared differences between the data and the predictions, as before. The initial values of the parameters are set as follows: the weights are all set to 1, scale values of the prior probabilities are set to their objective values, S_B is set to .8 and S_G is set to .2. Based on these initial estimates, the predictions are as shown in Figure 16.10. Insert Figure 16.10 about here.

Next, use the *Solver* to find the best-fit solution to the averaging model. The sum of squared discrepancies between data and theory is only 22.96 for this model, which is much better fit than for the Subjective Bayesian model. The solution is shown in Figure 16.11, and the predictions are graphed in Figure 16.12. Insert Figures 16.11 and 16.12 about here.

The averaging model provides a better fit to the data than either version of the Bayesian model. Apparently, people combine information by averaging the prior estimate with the evidence of the witness. When there is no witness, the relative weight of the prior information is greater than when there is a witness. This change in relative impact of information is evidence against the additive models and the subjective Bayesian models as well.

These results for the Cab Problem agree with the results and theory of Birnbaum and Mellers (1983) and Birnbaum and Stegner (1979). Birnbaum and Mellers (1983) asked judges to estimate the probability that a used car would last a certain period of time given the opinion of a mechanic who examined the car and base rate for that type of car. They also found that the effect of a source of information was inversely related to the number and credibility of other sources of information. These phenomena are evidence of averaging rather than additive models. Their study also investigated the effect of bias of witnesses, a factor not studied in the Cab problem.

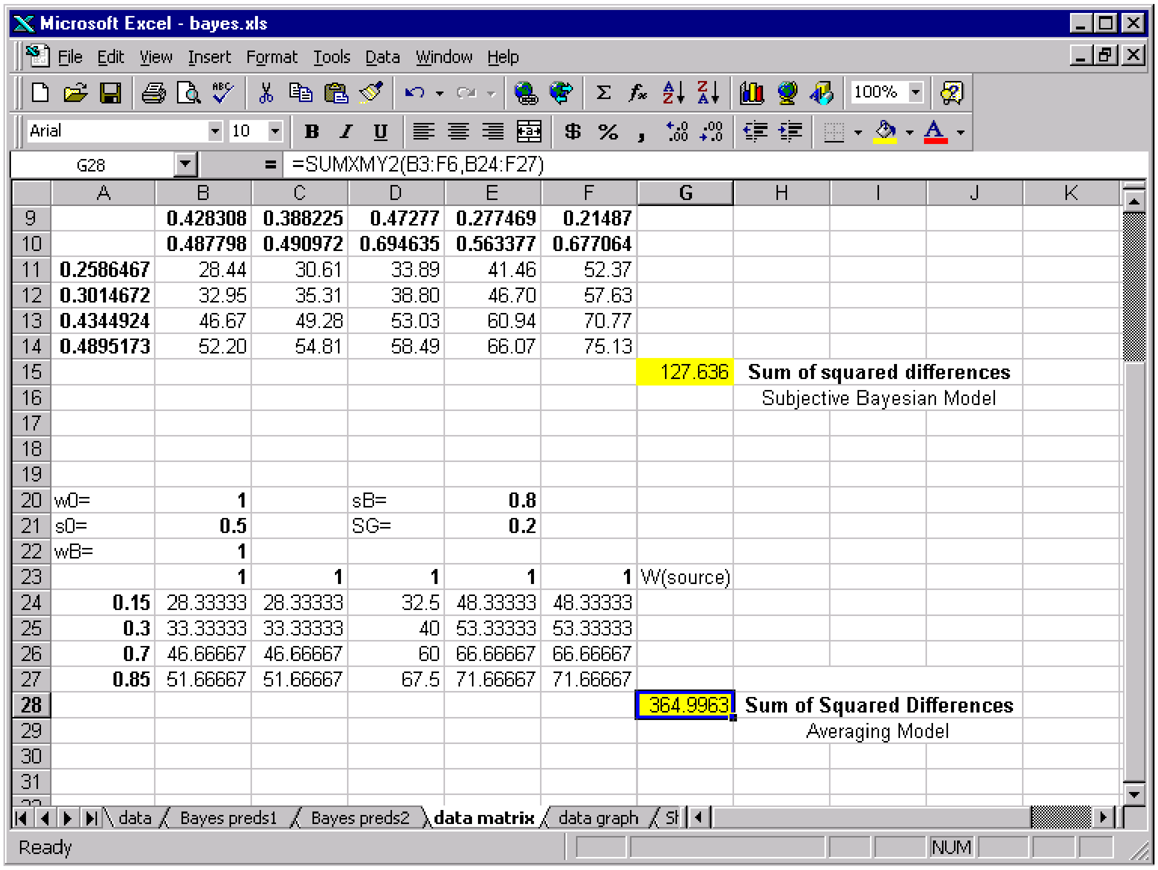
J. Summary

In this chapter you learned about Bayes theorem as a mathematical theory of how people should revise their beliefs and also as a descriptive theory attempting to predict how people do make inferences. The chapter reviewed a Web experiment, constructed with the help of surveyWiz, which replicated previous results obtained in the lab. People attend to the base rate, contrary to the claim of a base rate fallacy; however, people do not combine base rates with witness evidence by a Bayesian or additive process. Instead, they appear to average evidence. You should know how to construct such an experiment and analyze the results.

K. Exercises

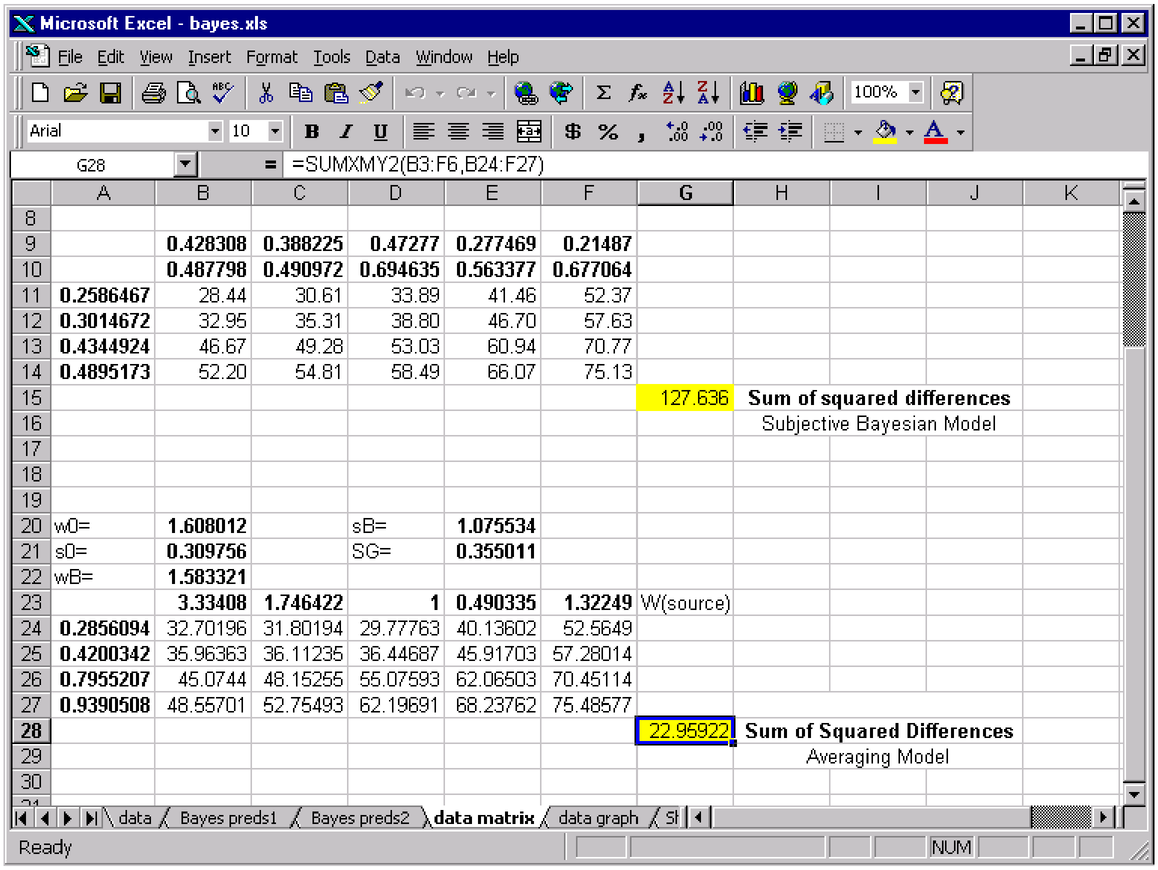
1. Examine the materials for the Cab problem in a text editor. Use factorWiz to create another variation of the Cab problem with different base rates: .20, .50, and .80. Use sources that have hit rates of .85 and false alarm rates of .15. This manipulation reverses numerical relationship between the witness and base rate; the source is now more diagnostic than the base rate. Also include a “source” of no witness.
2. Calculate predictions of Bayes theorem for the above 3 by 4 design using the Bayesian calculator.
3. Calculate predictions of Bayes theorem for the above 3 by 4 design using Excel.
4. Project idea: Construct the materials for a replication of the Cab problem study with the Lawyer/Engineer problem. In this problem, the judge is given the relative frequencies of lawyers and engineers in a sample, and is asked to infer if Tom is a lawyer or engineer based on a thumbnail description. For example, suppose Tom "likes jazz, and enjoys playing chess." What is the probability that Tom is a lawyer rather than an engineer? The proportion of lawyers or engineers is one factor (base rate), and the testimony and credibility of the thumbnail descriptions is the other factor.
5. Project idea: Construct the materials for an experiment on the bookbags and poker chips problem. An experimenter will draw a sample of chips from one of two bookbags. Bag R has 80 red chips and 20 white ones. Bag W has 20 red chips and 80 white ones. The experimenter decides randomly which bag to choose, and presents a random sample of chips from that bag to the judge. The judge's task is to infer which bag the sample came from. For example, suppose the experimenter flipped a coin, and chose one of the bags; suppose the sample of chips was 4 Reds and 1 White chip. Which bag do you think was chosen? What is the probability that the bag sampled was Bag R? There are many parameters that can be manipulated in this situation, making many interesting studies that can be done.

Figure 16.10. Spreadsheet set up to fit averaging model.



	A	B	C	D	E	F	G	H
9		**0.428308**	**0.388225**	**0.47277**	**0.277469**	**0.21487**		
10		**0.487798**	**0.490972**	**0.694635**	**0.563377**	**0.677064**		
11	**0.2586467**	28.44	30.61	33.89	41.46	52.37		
12	**0.3014672**	32.95	35.31	38.80	46.70	57.63		
13	**0.4344924**	46.67	49.28	53.03	60.94	70.77		
14	**0.4895173**	52.20	54.81	58.49	66.07	75.13		
15							127.636	**Sum of squared differences**
16								Subjective Bayesian Model
17								
18								
19								
20	w0=	**1**		sB=	**0.8**			
21	s0=	**0.5**		SG=	**0.2**			
22	wB=	**1**						
23		**1**	**1**	**1**	**1**	**1**	W(source)	
24	**0.15**	28.33333	28.33333	32.5	48.33333	48.33333		
25	**0.3**	33.33333	33.33333	40	53.33333	53.33333		
26	**0.7**	46.66667	46.66667	60	66.66667	66.66667		
27	**0.85**	51.66667	51.66667	67.5	71.66667	71.66667		
28							364.9963	**Sum of Squared Differences**
29								Averaging Model
30								
31								

Figure 16.11. Spreadsheet after *Solver* has found best-fit to the averaging model. Note that the sum of squared deviations for the Subjective Bayesian model are more than five times greater (worse) than the fit for the averaging model.


Microsoft Excel - bayes.xls

File Edit View Insert Format Tools Data Window Help

G28 = =SUMXMY2(B3:F6,B24:F27)

	A	B	C	D	E	F	G	H
8								
9		**0.428308**	**0.388225**	**0.47277**	**0.277469**	**0.21487**		
10		**0.487798**	**0.490972**	**0.694635**	**0.563377**	**0.677064**		
11	**0.2586467**	28.44	30.61	33.89	41.46	52.37		
12	**0.3014672**	32.95	35.31	38.80	46.70	57.63		
13	**0.4344924**	46.67	49.28	53.03	60.94	70.77		
14	**0.4895173**	52.20	54.81	58.49	66.07	75.13		
15							127.636	**Sum of squared differences**
16								Subjective Bayesian Model
17								
18								
19								
20	w0=	**1.608012**		sB=	**1.075534**			
21	s0=	**0.309756**		SG=	**0.355011**			
22	wB=	**1.583321**						
23		**3.33408**	**1.746422**	**1**	**0.490335**	**1.32249**	W(source)	
24	**0.2856094**	32.70196	31.80194	29.77763	40.13602	52.5649		
25	**0.4200342**	35.96363	36.11235	36.44687	45.91703	57.28014		
26	**0.7955207**	45.0744	48.15255	55.07593	62.06503	70.45114		
27	**0.9390508**	48.55701	52.75493	62.19691	68.23762	75.48577		
28							22.95922	**Sum of Squared Differences**
29								Averaging Model
30								

data / Bayes preds1 / Bayes preds2 / **data matrix** / data graph / SI

Ready NUM

Figure 16.12. Predictions of the Averaging Model. Note that this model correctly predicts the cross over of the *no witness* curve.