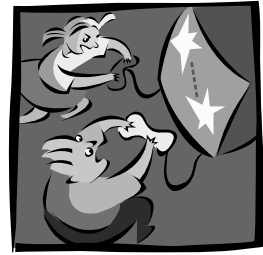**Gamz, Inc. Targets the Market**

Part I: Analyzing Information Drawn from a Marketing Database

GAMZ Inc. markets computer games online. It has a database of fifty thousand customers, aged 16 to 80. The database contains the recent purchase history of each customer, as well as three basic demographic data points: age, gender, and community type they live in – urban, suburban or rural. These customers have all purchased at least one game from GAMZ within the last 12 months, and no customer has purchased more than three games in the last 12 months.

GAMZ is planning on developing a high quality catalogue that describes a wide range of products targeted at different types of customers, called market segments. The fixed cost of developing the marketing campaign material is estimated to be $25,000. The variable cost of printing and mailing each catalogue is $2.50 per potential customer who receives a mailing. The computer games are priced at $90 each and the gross profit is $30 per game. The other $60 covers production costs and royalty payments to the developers but not mailing and catalogue costs.

1. If GAMZ carried out a mass mailing to the entire list, how many games would they have to sell to break even; that is, to recover the costs of developing and mailing the catalogue?

One statistic direct marketers use to decide which segments of a database to target is the response rate; that is, the proportion of customers contacted who bought the product.

2. What would the response rate have to be to break even?

3. What assumption did you make to determine this number?

4. Is this assumption realistic?

Upper management has said they are not really interested in any mass marketing effort unless it will generate at least $10,000 in profit.

5. How many games would they have to sell to make $10,000 profit?

GAMZ also uses a statistic called "profit value per customer" to account for the fact that some customers in the database have purchased more than one game in the past 12 months. Suppose the profit value per customer is $33.10 for the entire database.
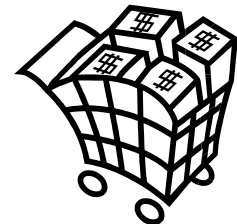
6. On average, how many games did each customer in the database purchase in the past year?

7. How does this information change the break even response rate?

8. Using the "profit value per customer" information, calculate the average response rate necessary to generate $10,000 in profit.

In order to maximize the response to a marketing promotion, marketers sometimes separate a customer database into various segments. A segment is a partition of the database that uses one or more factors such as gender, age, or location to characterize it. When a database has been partitioned into segments, every record in the database will belong to exactly one segment. GAMZ has decided to use gender and age to partition its customer database. GAMZ uses three categories of age: under 20, 20-29, and 30 and over. Thus, one of the segments it uses is "females under 20," and there are six segments altogether.

Table 1 contains data on the sizes, response rates, and profit for the last 12 months for the various segments of the GAMZ database.

| Segment | Size | Number of Buyers who Bought "*N*" games | | | Profit |
		N = 1	*N* = 2	*N* = 3	
Females under 20	9,029	8391	539	99	$292,980
Males under 20	9,407	8100	847	460	$335,220
Females, 20-29	8,202	7740	420	42	$261,180
Males, 20-29	9,221	8018	802	401	$324,750
Females 30 & over	7,789	7441	251	97	$247,020
Males 30 & over	6,352	6256	75	21	$194,070

Table 1: Response data for GAMZ customers.

9. For "Males under 20," how was the profit of $335,220 derived? That is, write an equation for that profit that uses the numbers in that row of the table.

    Optional: Enter the data from Table 1 into a spreadsheet, and translate your equation into a spreadsheet equation to calculate the profit for each segment.

10. Looking at the data in Table 1, the 9,029 females under 20 generated $292,980 in profit. The profit-value per customer of that segment of the market is therefore $292,980 /9,029 = $32.45. The standard deviations for the profit value per customer of each segment are provided. Compute the profit value per customer for each of the remaining market segments and enter those values in Table 2 below:

Segment	Size	Profit	Profit-Value	Standard Deviation
Females under 20	9029	$292,980	*$32.45*	$9.34
Males under 20	9407	$335,220		$15.01
Females, 20-29	8202	$261,180		$7.82
Males, 20-29	9221	$324,750		$14.41
Females 30 & over	7789	$247,020		$8.42
Males 30 & over	6352	$194,070		$4.71

Table 2: Profit-Value per Customer by Segment

11. Do any of the segments of the GAMZ database appear to be more or less promising as target segments for its marketing campaign? In your opinion, how great a difference in the profit value per customer would be of *practical* significance?

    Using a one-sided, two-sample Z-test, is the difference between the *largest* and *smallest* profit values per customer from Table 2 statistically significant? What is the value of p, the calculated level of statistical significance?

12. Using a one-sided, two-sample Z-test, is the difference between the *largest* and *next largest* profit values per customer from Table 2 statistically significant? What is the value of p?

13. If statistically significant, is either of these differences practically significant?

Repeat the two tests of significance from numbers 11 and 12 using the same means and standard deviations, but assume a sample size of 400 for each segment.

14. Is either difference statistically significant assuming the new sample sizes?   Explain.

News from the World of Operations Research

Merrill Lynch Applies Database Marketing Approach
Estimates Impact at $3.5-$6 Billion

Management Sciences at the Heart of OnStar's Success
GM Snares 80% Market Share

Merrill Lynch has developed a model that generates a "score" for potential clients. The scores are then used to prioritize the list of prospects. A cut-off score for qualifying prospects was established based on the probability that the potential customer would be a priority client.

Previous statistical studies at Merrill Lynch had yielded 85-90% accuracy in determining the characteristics of priority customers. These results suggested the potential of the database marketing approach for identifying priority customers from a database of prospects. Merrill Lynch conducted a field test of a newly selected set of prospects. The company estimates that the impact of the new approach will be between $3.5 and $6 billion in additional client assets.

Dbase Marketing at Fingerhut Reduces Catalogue Mailings
Profits Up $3.5 Million

The Mathematics of Decision-Making in Industry and Government
A Project of the Public Awareness Committee (PAC) of INFORMS

Project Coordinator:
  David Goldsman, Georgia Institute of Technology, Atlanta, GA
Series Editors:
  Kenneth R. Chelst, College of Engineering, Wayne State University, Detroit, MI
  Thomas G. Edwards, College of Education, Wayne State University, Detroit, MI
Operations Research Consultant:
  Jacob Zahavi, School of Business, Tel Aviv University, Tel Aviv, Israel

Teacher Writing Team:
  Marlene Lawson, Maggie L. Walker
    Governor's School for Government &
    International Studies, Richmond, VA
  Hazel Orth, Langley H.S., McLean, VA
    (retired)

Part II: Using Statistical Analysis to Make a Marketing Decision

In order to increase the response to its promotion, the Marketing Department at GAMZ, Inc. is considering using a $10-off incentive on the initial offer. They decide to test the response to the promotion with or without the incentive. To do so, they draw two random samples of 1,000 customers each from their 50,000-customer database. One of these test samples receives the promotion catalogue with the $10-off incentive. The other test sample receives the promotion catalogue only, with no additional incentive. Recall from part I that the profit per game is $30, and the cost of printing and mailing is $2.50 per potential customer.

1. For the test sample that receives the incentive, what is the net profit per game sold?

2. Using the net profit per game sold for the test sample that receives the incentive, calculate the response rates needed to break even and to return at least a $10,000 profit, as you did in Part I, questions 7 and 8.

Table 3 contains a summary of the test data for the two samples.

With Incentive:

Segment	Total sample size	Number of responses	Number of responders who bought: 1 game	2 games	3 games	Total games purchased	Response rate	Net profit per Segment
Females under 20	181	27	23	3	1	32	0.149	
Males under 20	191	32	25	5	2	41	0.168	
Females 20-29	158	16	14	2	0	18	0.101	
Males 20-29	193	25	18	4	3	35	0.130	
Females 30 & over	132	15	14	0	1	17	0.114	
Males 30 & over	145	18	17	1	0	19	0.124	
TOTAL	1000	133	111	15	7	162	0.133	

Without Incentive:

Segment	Total sample size	Number of responses	Number of responders who bought: 1 game	2 games	3 games	Total games purchased	Response rate	Net profit per segment
Females under 20	175	16	13	2	1	20	0.091	
Males under 20	184	19	14	2	3	27	0.103	
Females 20-29	165	23	21	2	0	25	0.139	
Males 20-29	180	24	19	4	1	30	0.133	
Females 30 & over	150	16	14	2	0	18	0.107	
Males 30 & over	146	11	11	0	0	11	0.075	
TOTAL	1000	109	92	12	5	131	0.109	

Table 3: Sample Test Data With and Without an Incentive

3. Calculate the net profit per segment for both test samples. Do any of the segments of the GAMZ database appear to be more or less promising as target segments for the marketing campaign?

4. Construct a 90% confidence interval for the difference between the response rates for the two test groups.

5. Use the data in Table 3 to state null and alternate hypotheses regarding the difference between the response rate of the test sample that received the incentive and the response rate of the test sample that did not receive the incentive.

   Perform a one-sided two-proportion Z-test of your hypothesis at the 95% confidence level, and state the results of the test.

6. Ignoring the cost of producing and mailing the catalogue, compute the total net profit for each test sample.

7. Looking back at the response rates that you calculated in questions 7 and 8 in Part I, perform a one-sided one-proportion Z-test at the 90% level of confidence of the following hypotheses if the incentive is *not* used:

   a. H_0: The response rate will not be sufficient to break even.
      H_a: The response rate will be sufficient to break even.

      If the incentive is *not* used, is the promotion likely to break even?   What level of confidence do you have in your response?

   b. H_0: The response rate will not be sufficient to earn at least $10,000 in profit.
      H_a: The response rate will be sufficient to earn at least $10,000 in profit.

      If the incentive is *not* used, is the promotion likely to generate $10,000 in profit?   What level of confidence do you have in your response?

8. Looking back at the response rates that you calculated in question 2 in Part II, perform a one-sided one-proportion Z-test at the 90% level of confidence of the following hypotheses if the incentive *is* used:

   a. H_0: The response rate will not be sufficient to break even.
      H_a: The response rate will be sufficient to break even.

      If the incentive *is* used, is the promotion likely to break even?   What level of confidence do you have in your response?

   b. H_0: The response rate will not be sufficient to earn at least $10,000 in profit.
      H_a: The response rate will be sufficient to earn at least $10,000 in profit.

      If the incentive *is* used, is the promotion likely to generate $10,000 in profit?   What level of confidence do you have in your response?

The Marketing Department at GAMZ, Inc. has decided to go forward with the promotion, but not to offer the $10-off incentive.

9. Explain why this decision does or does not make sense, based on your tests of the available data.

Next, the Marketing Department decides to investigate the test sample that did not receive the incentive to learn whether the response rates for any segments of the database were higher than would be expected. If so, it would make sense to target such segments in the promotion. Table 4 contains the observed numbers of responders and non-responders for each segment of the test sample that did not receive the incentive.

Observed	Females under 20	Males under 20	Females 20-29	Males 20-29	Females 30 & over	Males 30 & over	Totals
Responders	16	19	23	24	16	11	109
Non-Responders	159	165	142	156	134	135	891
Totals	175	184	165	180	150	146	1000

Table 4: Observed Numbers of Responders and Non-Responders by Segment

Notice that of the 1000 customers in the sample who did not receive the incentive, there were 175 females under 20. This is 0.175 of the sample. If there is no segment of the database that is more likely to respond than any other, the expected value of the number of responders for females under 20 is 0.175 of the 109 total responses; that is, (0.175)(109)=19.075. Similarly, the expected value of the number of non-responders is 175–19.075=155.925, as shown in Table 5.

10. Compute the expected value of the number of responders and non-responders from each of the other five segments of the database and enter those values in Table 5.

Expected	Females under 20	Males under 20	Females 20-29	Males 20-29	Females 30 & over	Males 30 & over	Totals
Responders	19.075						
Non-Responders	155.925						
Totals	175						

Table 5: Expected Numbers of Responders and Non-Responders by Segment

Now, a Chi-square test of association may be performed to test the null hypothesis that the observed response rates across all of the segments are the same against the alternate hypothesis that these response rates differ significantly. To calculate the value of Chi-square, determine for each segment the difference between the observed number of responders in Table 4 and the expected value of that number in Table 5. Similarly, the difference between the observed number of non-responders and its expected value must be computed. Those differences are then squared, and the squared differences are divided by the expected values. Finally, these twelve terms are summed to get the value of Chi-square.

The mathematical formula is: $\chi^2 = \sum \frac{(o-e)^2}{e}$, where

o and *e* represent the observed numbers and their expected values, respectively. Obviously, computing the value of Chi-square in this case involves a good deal of tedious computation.
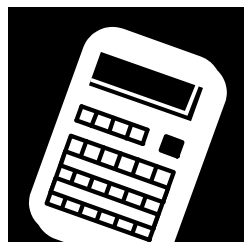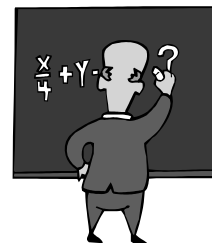
Fortunately, statistics software on computers or calculators can be utilized to accomplish this task. The set-up for computing Chi-square using a TI-83 Plus calculator is described below. To use a different calculator or computer software, refer to the documentation accompanying the calculator or software.

To begin the set-up, the information in Table 4 above must be stored in a 2x6 matrix, *A*, on the calculator. To do so, press "2nd" and "MATRIX" to obtain the menu for working with matrices, highlight "1:" and "EDIT", and press "ENTER". You may now edit "MATRIX [A]". Notice that the default dimensions are 1x1, so change that to 2x6 and begin entering the values from Table 4. Matrix *A* is too wide to be viewed on one screen, so the last two screens in the set-up show what the completed matrix *A* should look like.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

Figure 1: TI-83 Matrix Set-Up

Now, to calculate the value of Chi-square, press the "STAT" key to access the statistics menu, highlight "TESTS" and "C:" ("χ^2-TEST …"), and press "ENTER". Then highlight "CALCULATE" and press "ENTER" again.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompresso
are needed to see this picture.

Figure 2: Calculating the Value of Chi-square Using a TI-83

11. Compute the value of χ^2.

12. For these data, what is the probability of obtaining a value of Chi-square at least as large as the computed value?

13. From the output of the computation of the value of Chi-square, how many degrees of freedom did this computation have?

One way to interpret the number of degrees of freedom in a data set is as the number of *independent* pieces of information contained in the data set. If any piece of information in the table may be computed from other pieces of information in the table, then it is not an independent piece of information.

13. In light of the above interpretation of degrees of freedom, explain why your answer to question 12 does or does not make sense.

14. Would you recommend to the marketing Department at GAMZ, Inc. that they target specific segments of their database in the promotion?   Why or why not?

Finally, the Marketing Department at GAMZ, Inc. must forecast the total expected net profits from the promotion in order to estimate whether the promotion is viable. Recall that management requires a net profit of at least $10,000 to go forward with any promotion. If the promotion is mailed to the remaining 48,000 customers in the database, use the data from the test sample that did not receive the incentive to answer the following questions.

15. What is the expected value of the number of purchasers?

16. On average, what is the forecasted total number of games that will be purchased?

17. Compute the expected value of the gross profit on the sale of the games.

Recall from Part I that there is a fixed cost of $25,000 associated with the promotion, as well as a cost of $2.50 for each customer to whom the promotion is mailed.

18. Compute the expected value of the total net profit on the entire promotion, excluding the two test samples.

Does it appear that the promotion will meet management's requirement of net profits of at least $10,000?