
Abstract
Advertising plays a key role in almost every firm's marketing strategy. This paper investigates two relatively unexplored areas of advertising: (1) the optimal amount of information an ad should contain, and (2) what information an ad should contain. An application to point-of-purchase advertising (store displays) is given.

Introduction
Advertising plays a key role in almost every firm's marketing mix. Point-of-purchase advertising, also known as store displays, are one of the retailer's "most powerful sales promotion forces" (Davidson and Doody, 1966). Consequently, the literature (e.g., Aaker and Myers, 1975; Roman and Maas, 1976) on advertising continues to progress in many directions. Models have been constructed to determine optimal advertising expenditures (Nerlove and Arrow, 1962; Borey, 1977), optimal amount of copy testing (Gross, 1972), and proper media allocation (Little and Lodish, 1969; Gensrich, 1973). However, despite pioneering studies discussing advertising strategies (Boyd et al., 1972; Nelson, 1974), many important aspects of advertising copy use are left unexplored. For example, to this author's knowledge no model exists that optimizes both the amount and nature of information the ad should contain.

Why do consumers read or view advertisements? Well, they must expect some benefit. Perhaps they receive some information. But what information do these consumers receive and how does it affect consumer choice? The purpose of this paper is to address these important issues. The paper first discusses an objective of advertising, i.e., seduction. Once seduction is understood, the paper continues to show how to maximize the probability the ad will seduce the user of a competitive brand. Finally, the paper concludes with an illustrative example dealing with the design of point-of-purchase advertising.

A Theory of Seduction
The Process of Seduction
When someone attempts to persuade someone else to purchase some merchandise, either at the point of purchase or before the point of purchase, the mechanism used can be construed as a process of providing information. The information can be concrete, factual, and very straightforward. For example, in a cigarette ad an advertiser may state that their brand contains only 5 mg. of tar. The informed reader or viewer of this ad now has more information. This information may persuade the reader to switch to the advertised brand. The information can also be abstract, subjective, and less obvious. For example, a pantyhose ad portrays a woman using the advertiser's brand at a party. The other women at the party are using competitive brands or have legs that are not visible. From the positions of those involved, and the interest of the male participants, the message of the ad becomes clear. The informed reader of the ad now is aware of a hidden quality of the product; that is, the product's ability to attract members of the male gender. In any case, whether information is hidden or not, the ad attempts to provide information that will tend to favorably influence the reader's choice process.

However, there is a difference between enlightenment and seduction. With enlightenment, the consumer (decisionmaker) is given information to enable better decisionmaking. With seduction, information must be provided so as to have a desirable and predictable outcome on the consumer's decision process. Clearly, seduction may not be easy. First, consumers may realize that advertisers have the incentive to provide biased information. Second, consumers may have very little incentive to read or view advertisements and may thereby circumvent enticement completely. We therefore require some understanding of the consumer choice process, including information acquisition and information processing (Simon and Newell, 1971; Bettman, 1977; Lehmann, 1977), before we can determine the optimal enticing strategy.

Let us assume that consumer choices are determined directly or as a cue by the product's characteristics as defined by Lancaster (1966) and the product's price. Then, a consumer who wishes to choose between two products may proceed by comparing the two products on their characteristics. For example, a consumer choosing between two household cleaning products may first compare them on ammonia content. Second, a comparison on drying speed may take place. Next, the products could be compared on the attractiveness of their respective colors. These comparisons could then proceed until all characteristics of the product are exhausted, uniquely defining each product. Undoubtedly, this process can be lengthy. The two products may be compared a multitude of characteristics (possibly infinite) before choosing, with certainty, the product that is most preferred. The process can, therefore, be rather tiresome and likely to end before all characteristics are exhausted.

Of course, this representation of the consumer-choice process would be void of implications without a methodology to describe and classify choice situations with respect to the number of necessary comparisons in order to reach a decision. Fortunately, this representation can be interpreted as a sampling problem (Wald, 1947; Mood et al., 1974). The consumer can be viewed as sampling product-pair differences by characteristic. For example, consider the household cleaning product comparison. The consumer first compares the products on ammonia content. This comparison is basically sampling from the population of product differences. The first observation chosen was difference in ammonia content. Again, the second comparison on drying speed can be viewed as an observation on drying-speed difference. It is then possible to show, given some unrestricted assumptions, that the consumer must balance expected benefits against decisionmaking costs. As a result, each consumer will seek some N* amount of information about a product before making a decision. Even if a product has been previously purchased, many times the consumer may still seek a minimal amount of information to reinforce existing beliefs and possibly seek disconfirming evidence to enable better future decisions. We would, then, expect to find a heterogeneous population with each member conceptually having a different value for N* . We will now explore a specific implication of this descriptive theory toward explaining seduction. General implications, in addition to the strengths and weaknesses of the theory, are discussed elsewhere (Shugan, 1980).
The Advertiser's Seduction Problem

In reality, we seldom find companies whose brands dominate their competitors on every characteristic. Hence, it would not be prudent for most advertisers to reveal to the consumer every last detail of their brand. As a result, enticing the consumer to purchase a brand would seldom involve the revealing of the characteristics on which the brand is weak. The retailer or advertiser must be concerned with exactly how much and what, if any, information should be provided to the potential buyer through the ad.

Let us begin by looking at only that segment of the market that can be seduced by the advertisement. This segment is defined so that if any member of the segment was exposed to all the brand's favorable characteristics, that member would purchase the brand. The advertiser need only consider this segment because, by definition, the advertisement cannot entice others to buy despite the effectiveness of the ad.

Each member of the seduceable segment seeks information about some nong啷egeous number of product characteristics. Define P(n*) for n* > 0 as the proportion of seduceable individuals who are seeking information on n* product characteristics. Then P(n*) represents the probability of a randomly chosen member of the segment desiring information on n* characteristics. Consequently, if an ad (or display) contains information on N characteristics, individuals seeking n* < N information will sample the ad rather than spend the time and effort to read the ad in its entirety. Individuals with n* ≥ N will read the ad in its entirety and may, if left unconvinced, continue to seek information about the product. Let y be the probability an individual will look, or be exposed to, any one characteristic in the ad given the ad contains information on N characteristics. This probability can be computed by using equation (1).

\[ L_N = \sum_{X=0}^{N} \left( \frac{n^*}{N} \right) P(n^* = X) \left( \frac{N}{X} \right) \]

where \( P(n^* = X) \) is the probability a randomly selected individual is seeking information on X characteristics. Now, we wish to maximize the probability a seduceable individual is seduced. That probability can now be expressed as shown in equation (2).

\[ S_N = \sum_{j=1}^{N} E_j N_j \]

where \( S_N \) is the probability of seducing a randomly chosen individual, \( N_j \) is the probability the individual will perceive any one characteristic as given by equation (1), and \( E_{N_j} \) is the probability the individual will be excited to purchase the product by characteristic \( j \).

To make the theory more concrete, make two simplifying assumptions, which are later relaxed. Assume: (1) If the advertiser listed all of the brand's characteristics in order of their convincing power (i.e., the superiority over competitive products), the probability of a particular characteristic convincing a randomly chosen member of the population decreases at a constant rate \( \lambda \) as we go down the list. (2) The proportion of people seeking information on n* characteristics decreases at a constant rate \( r \) as n* increases.

Finally, assume \( E_{N_j} \) is a function of the number of characteristics in the ad. This, for example, assumes no synergism (the convincing power of two characteristics is the sum of their respective powers). Assumptions (1) and (2) lead to equations (3) and (4) (e.g., see Mood et al., 1974).

\[ P(n^* = X) = r(1 - r)^X \]  
\[ E_{j} = \lambda(1 - \lambda)^j \]

Recall that the characteristics are ordered from \( j = 1 \) to \( j = \infty \) in terms of their convincing power. Later, measurement of these parameters will be discussed.

In order to make the following development less awkward and the optimization less clumsy, the discrete geometric distributions (equations (3) and (4)) for \( P(n^*) \) and \( E_j \) will be replaced by their continuous analogues, i.e., exponential distributions. The continuous analogue to the discrete problem is then given by equation (5).

\[ S_N = \int_{0}^{\infty} \left( \frac{n^*}{N} \right) e^{-n^*y} \left( \frac{N}{n^*} \right) \lambda e^{-\lambda y} dy \]

To find the optimal amount of information the advertiser should provide, we must maximize \( S_N \) with respect to \( N \). Then, using Leibnitz's rule for differentiating an integral—integrating by parts, setting \( dS_N/dN \) equal to zero, and rearranging—yields equation (6).

\[ \frac{1}{N^*} = \left( \frac{1}{e^{\lambda r} - 1} \right) + \left( \frac{\lambda}{e^{\lambda r} - 1} \right) \]

where \( N^* \) is the optimal amount of ad information. This equation has no nice closed-form solution for \( N^* \), and must be solved iteratively for \( N^* \).

Using the iterative method shown in the appendix, Table 1 was generated. This table shows the optimal ad information content \( N^* \) as a function of \( r \) and \( \lambda \).

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( r = .1 )</th>
<th>( r = .3 )</th>
<th>( r = .5 )</th>
<th>( r = .7 )</th>
<th>( r = .9 )</th>
<th>( r = 1.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1</td>
<td>12.5</td>
<td>7.5</td>
<td>5.7</td>
<td>4.9</td>
<td>4.0</td>
<td>3.3</td>
</tr>
<tr>
<td>.3</td>
<td>6.7</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>.5</td>
<td>4.8</td>
<td>5.1</td>
<td>5.2</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>.7</td>
<td>3.7</td>
<td>4.2</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>.9</td>
<td>2.9</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
</tr>
<tr>
<td>1.0</td>
<td>2.6</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>1.1</td>
<td>2.4</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>1.2</td>
<td>2.2</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
</tr>
</tbody>
</table>

We find \( N^* \) provides the proper balance between revealing too much about the brand and not revealing enough. If the advertisement does not reveal a sufficient number of the
brand's attractive characteristics, the ad viewer may remain unconvincing and not purchase the brand. The ad would not be sufficiently revealing to entice the consumer. If the advertisement reveals too many of the brand's characteristics, the viewer may not notice all of the brand characteristics. In this case, there is the danger the viewer may focus on some of the brand's less attractive characteristics and ignore some of the more convincing characteristics.

Some Qualitative Implications

The probability a consumer will buy our brand with exposure 1 to less than one characteristic is \( \int_0^1 \exp(-\lambda x) dx \) or 1 - \( \exp(-\lambda) \). However, this quantity represents our current market share computed as a percentage of our total potential market share. Hence, if we let \( m \) denote our current share, then \( \lambda = -\log(1 - m) \). Consequently, new products or products that are repositioned will have smaller \( \lambda \) (because \( m \) is smaller). Therefore, new product ads should contain relatively more information than mature product ads. Further, brands that have been well managed require less ad information than brands that are far from their potential.

Similar analysis applies to \( r \). In this case, we find that ads should contain relatively more information when products show greater variability on important characteristics, products serve different markets, product prices are high relative to total income, and consumers have very different tastes (Shugan, 1980).

Measuring the Relevant Parameters

For the advertiser or retailer to actually apply the preceding analysis and solve for the optimal \( N^* \), that advertiser must ascertain how to measure the parameters \( r \) and \( \lambda \). It would also be convenient if the advertiser could determine these parameters using existing market research techniques and possibly from existing data bases. We will, therefore, try to find measurements of \( r \) and \( \lambda \) available from established techniques rather than develop new untested measurement instruments. However, better measurement methodology may be a desirable goal for future research.

Consider for a moment consumers who tend to look at only a few product characteristics. These consumers may infer the values of other characteristics based on what they find for the few characteristics they observe. Suppose we ask these consumers to state their perceptions by rating a familiar brand. If we then analyze the consumers' perceptions to develop a model of these perceptions, we should find the majority of their perceptions should be highly correlated and explained by only a few underlying factors. Further, if we ask a respondent \( n^* \) perceptual rating questions and assume a linear relationship between \( n^* \) and the respondent's perceptual variability for characteristic \( j \), then equation (7) expresses the distribution of response variability.

\[
f(x_j^2 | x_j; \nu, \tau) = \begin{cases} 
\frac{\tau \nu^{X/V}}{\Gamma(X/2)\Gamma(V/2)} \exp\left(-\frac{x_j^2 + \tau^2 V}{2 \nu} \right) & \text{for } X < V \\
\frac{\tau \nu^{V/2}}{\Gamma(V/2)} & \text{for } X = V \\
0 & \text{for } X > V
\end{cases} \tag{7}
\]

where \( \nu^2 \) is the variability in respondent \( j \)'s answers and \( \nu \) is the true characteristic variability. The maximum-likelihood estimator for \( \tau \), denoted \( \hat{\tau} \), is given by equation (8).

\[
\hat{\tau} = J \left[ \frac{1}{(n/V)} \sum_{j=1}^{J} \sigma_j^2 \right] \tag{8}
\]

where \( V = \text{maximum } \sigma_j^2 \) and \( J \) is the number of respondents.

Interestingly, this procedure is related to the measurement of halo effects (Thorndike, 1920; Beckwith and Lehmann, 1975). Now, given a method for estimating \( r \), we require a method for estimating \( \lambda \). We have postulated that an individual is convinced to buy one brand over another by sampling product characteristic differences. That is, the individual would have observations \( z_j \) where \( z_j \) is defined as follows:

\[
z_j = U_{A,j} - U_{B,j}
\]

where \( U_{A,j} \) is the utility of characteristic \( j \) for brand \( A \), and \( U_{B,j} \) is the utility of characteristic \( j \) for brand \( B \).

There exists intensity measures for \( U_{A,j} \) (Hauser and Shugan, 1980). Unfortunately, most other more well-known market research techniques, such as conjoint analysis (Luce and Tukey, 1964; Green and Srinivasan, 1978) and direct assessment (Hauser and Urban, 1979), yield functions that are only unique to a positive linear transformation. However, we note:

\[
\lambda = \exp\left(\frac{-X_j z_j}{\nu^2} \right) \tag{9}
\]

and, therefore, because

\[
\lambda = \frac{\log(z_j / z_{j+1})}{X_j} \tag{9}
\]

we can use \( (z_j / z_{j+1}) \) to estimate \( r \). Then, one estimator for \( \lambda \) that is robust to the arbitrary scaling for \( U_{A,j} \) is given by equation (10):

\[
\hat{\lambda} = \frac{1}{N-1} \sum_{j=1}^{J} \log(z_j / z_{j+1}) \tag{10}
\]

which minimizes the mean squared error in equation (9).

Finally, before proceeding, we should note that no attempt has been made to find the distribution of these estimators. It would, therefore, be inappropriate to attempt to do any statistical testing.

Now that we have a method for measuring both \( r \) and \( \lambda \), as given by equations (8) and (10), we can proceed to an illustrative application of the theory.

Illustrative Example

The model developed in the preceding analysis can be used to either design traditional advertisements or help a retailer construct more affective product displays. One major supermarket wanted to encourage customers to switch from lower-margin brands to higher-margin brands. The following example illustrates how the optimal display can be determined. Respondent data are simulated and the example is hypothetical. However, the example parallels analysis currently being performed.

Let us consider three product categories: eggs, coffee, and soup. A focus group (see Alpert (1971) for alternatives) was used to determine the relevant discriminating product characteristics. Table 2 lists the characteristics identified by focus groups for each product category.
(7) Take the competitor's brand we wish to position our brand against. Rank the characteristics according to the magnitude of the differences determined in step (6). Table 4 shows these ranked differences for our brand's best four characteristics. For example, our soup's best competitive characteristic is our super ingredients when competing with brand 2 (a difference of 1.4); our second best competitive characteristic is our extra-large can size (a difference of 1.3).

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>DIFFERENCE IN UTILITY WITH OUR BRAND ON CHARACTERISTIC j (ordered from largest Z_j to smallest Z_j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggs</td>
<td>Brand 1: Z_j = 1.7 Z_1 = .9 Z_3 = .1 Z_14 = .1</td>
</tr>
<tr>
<td>Soup</td>
<td>Brand 1: Z_j = 1.7 Z_4 = 1.3 Z_3 = .7 Z_10 = .2</td>
</tr>
<tr>
<td>Soup</td>
<td>Brand 2: Z_j = 1.4 Z_3 = 1.3 Z_10 = .6 Z_4 = .1</td>
</tr>
<tr>
<td>Soup</td>
<td>Brand 3: Z_4 = 1.1 Z_7 = .8 Z_3 = .4 Z_10 = .2</td>
</tr>
<tr>
<td>Coffee</td>
<td>Brand 1: Z_7 = 1.6 Z_9 = 1.1 Z_2 = .8 Z_12 = .3</td>
</tr>
<tr>
<td>Coffee</td>
<td>Brand 2: Z_7 = 1.7 Z_9 = .9 Z_18 = .7 Z_3 = .7</td>
</tr>
</tbody>
</table>

(8) From data found in step (7), use equation (10) to estimate λ. Results are summarized in Table 5.

(9) Compute the variance of the data in step (1) and use equation (8) to estimate r. Estimates for r are shown in Table 5.

<table>
<thead>
<tr>
<th>TABLE 5</th>
<th>ESTIMATED r AND λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggs</td>
<td></td>
</tr>
<tr>
<td>Estimated r</td>
<td>.91</td>
</tr>
<tr>
<td>Estimated λ</td>
<td></td>
</tr>
<tr>
<td>Brand 1</td>
<td>.83</td>
</tr>
<tr>
<td>Brand 2</td>
<td>--</td>
</tr>
<tr>
<td>Brand 3</td>
<td>--</td>
</tr>
</tbody>
</table>

(10) From steps (8) and (9), use Table 1 (or equation (6)) to find the optimal ad information content. Results are summarized in Table 6.

<table>
<thead>
<tr>
<th>TABLE 6</th>
<th>OPTIMAL AD INFORMATION CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggs</td>
<td>Soup</td>
</tr>
<tr>
<td>Brand 1</td>
<td>1.4</td>
</tr>
<tr>
<td>Brand 2</td>
<td>--</td>
</tr>
<tr>
<td>Brand 3</td>
<td>--</td>
</tr>
</tbody>
</table>

After determining the product characteristics consumers use for discriminating among brands, the following methodology can be employed.

Methodology

(1) Determine how the consumer perceives each competitive brand on each characteristic. In this example, agree/disagree questions were used, as illustrated in Table 3.

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>SAMPLE AGREE/DISAGREE QUESTIONS FOR MEASURING PERCEPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Brand A eggs provide the highest possible quality.</td>
<td>Strongly Agree</td>
</tr>
<tr>
<td>2. Brand B soup is too salty.</td>
<td>a</td>
</tr>
<tr>
<td>3. Brand A coffee has a bitter taste.</td>
<td>a</td>
</tr>
</tbody>
</table>

(2) Determine how the consumer perceives our brand on each characteristic.

(3) Determine the utility the consumer derives from each characteristic at each level. For example, preference regression (Urban and Hauser, 1980), conjoint analysis (Luce and Tukey, 1964; Green and Srinivasan, 1978), or intensity measures (Hauser and Shugan, 1980) could be used.

(4) From steps (2) and (3), calculate the utility the consumer derives from each of our brand’s characteristics.

(5) From steps (1) and (3), calculate the utility the consumer derives from each competitive brand’s characteristics.

(6) From steps (4) and (5), find the difference in utility between our brand and each competitive brand on each characteristic.
Discussion

From the preceding analysis, we see that r, and, hence, the optimal ad information, depends on our product positioning. Depending on what brand we position against, our optimal advertising strategy varies. The positioning decision, however, depends on a host of factors whose discussion is beyond the scope of this paper. Other sources exist on optimal product positioning (Striefer, 1968; Hauser and Urban, 1973; Silk and Urban, 1978; Hauser and Urban, 1979; Hauser and Shugan, 1980; Urban and Hauser, 1980). We will, therefore, take product positioning as given.

In this example, the positioning strategy depends on the relative profit margins for the brands involved. Table 6 implies the optimal display strategy for each positioning strategy. For example, if our coffee is positioned against Brand 1, then our display should stress three of our product's characteristics: 7—our super ingredients; 4—our chunkiness; and 3—our extra ounce (can size). If, however, we seek to position our soup against Brand 2, our display should discuss only 2-1/2 of our product's characteristics. Perhaps 4/5 of our display (2/2.5) should be devoted to characteristics 7 and 3, while 1/5 of our display (.5/2.5) should be devoted to characteristic 10—i.e., our meat content. Similar strategies are implied for coffee and eggs.

It is important to remember this discussion says nothing concerning display size or the effectiveness of its presentation. The former problem involves optimal advertising budgets, and the latter problem requires copy testing.

Summary and Conclusions

This paper has made an attempt to rigorously examine the problem of optimal ad information content. We found the advertiser must tradeoff too much information content against too little. By making some simple assumptions, the theory became operational. An illustrative example showed how the theory could be applied in practice to help solve one important problem for retailers—i.e., display design. Of course, the model presented was only a first step. However, it demonstrated that complex and apparently extremely qualitative factors, such as the difficulty of decisionmaking, the effort expended in reading an ad, the amount of information content in an ad, and the seductive power of an ad, were not beyond the grasp of rigorous analysis.

Much work still needs to be done. Better measurements must be found. Effectiveness of the communication needs to be explored. Multibrand positioning problems must be attacked. A system for defining characteristics must be developed. The model must be extended so to allow one characteristic to provide information about other characteristics. Finally, competitive actions should be considered.

Appendix

Algorithm to Determine N* for e^x. Performing this substitution and rearranging, we find the optimal N* must satisfy equation (12).

\[ (\ln N/2!) + (\ln N/3!) + ...] \ln (\lambda N/2!) + \ln (\lambda N/3!) + ... = 1 \]

where x! = 1 * 2 * 3 ... x. Equation (12) reveals that the value of N* is unique because both terms in equation (12) are monotonic in N. Moreover, equation (12) allows us to find approximate solutions for N*. For example, a first-order approximation solution for N* is given by (rN*/2!)(XN*/2!) - 1, which implies

\[ N* = 2/\sqrt{x} \]

The expression shown in equation (13) provides a starting solution for N* when solving the problem iteratively.

To obtain a solution for N within E/2 of the true N, use the following algorithm. Let f(x) = [r/(e^x - 1)] + [\lambda/(e^x - 1)] - 1/N.

Start

Set L = 1/\sqrt{x}
Set U = (1/r) + (1/\lambda)
1

N = (L + U)/2

[U - L] < c

f(N) < 0

no

no

Set L = N

Set U = N

END
Despite the convenience of assumptions (1) and (2) in the text, for some situations these assumptions may be unreasonable. For example, the advertiser may feel the brand's best characteristics are equally convincing. In this case, a uniform distribution should be used rather than an exponential distribution. In the general case, we can assume:

(1) If the advertiser listed all of the brand's characteristics in order of their convincing power, the probability at the $x_{ft}$ characteristic convincing a randomly chosen member of the population is described by the density function $f(x)$.

(2) The proportion of people seeking information on $x$ characteristics is described by the density function $g(x)$.

Then the probability of seducing any given member of the seducible segment is given by equation (14).

$$\sum_{n=0}^{N} \int_{0}^{X} f(x) dx = g(x) \int_{0}^{N}$$

The advertiser wishes to maximize $S_N$ with respect to $N$. Therefore, we set $dS_N/dN = 0$. Define $q(x), f(x)$, and $G(x)$ so that:

$$d^2q(x) = 4F(x) = f(x)$$

$$dG(x) = g(x)$$

Then, using Leibnitz's rule, integrating by parts, rearranging, and setting $dS_N/dN = 0$, we find the optimal $N^*$ must satisfy equation (17).

$$\frac{N^2 P(N^*) - [q(N^*) - q(0)]}{N^2 [1 + F(0)] - [q(N^*) - q(0)]} = \frac{g(N^*)}{G(N^*)}$$

If $N^*$ satisfies this equation, then $N^*$ will maximize the seductive power (i.e., probability of seduction) of the advertisement. Intuitively, we find $N^*$ provides the proper balance between revealing too much about the brand and not revealing enough. If the advertisement does not reveal a sufficient number of the brand's attractive characteristics, the ad viewer may remain unconvinced and not purchase the brand. The ad would not be sufficiently revealing to entice the consumer. If the advertisement reveals too many of the brand's characteristics, the viewer may not notice all of the brand characteristics. In this latter situation, the consumer samples the ad for product characteristics and, in doing so, may infer the brand as being less attractive than in the situation where the ad only contained the more convincing attributes.

Maximum-Likelihood Estimators for $f$

If $\sigma^2$ and $n^*$ are linearly related, then for some $\alpha$ and $\beta$, $\sigma^2 = \alpha + \beta n^*$. Further, $\alpha = 0$ and $\beta = V/M$ because $V > 0$ and $\alpha = 0 + 80$ when $n^*$ equals $M$ and $0$, respectively. (Remember, $V$ is the actual characteristic variability in the $N$ attributes in the perceptual rating questions.) Then:

$$f(0^2 = X; r, V) = \begin{cases} \frac{P[(V/M)n^* = X]}{\sum_{n^*} P[n^* = N]} & \text{for } X < V \\ \frac{P[n^* = N]}{\sum_{n^*} P[n^* = N]} & \text{for } X = V \\ 0 & \text{for } X > V \end{cases}$$

and

$$f(0^2 = X; r, V) = \begin{cases} \frac{-e^{-X(V/M)}}{\sigma} & \text{for } X < V \\ \frac{e^{-N/V}}{\sigma} & \text{for } X = V \\ 0 & \text{for } X > V \end{cases}$$

so we maximize

$$\sum_{j=1}^{J} \frac{f(0^2 = X_j; r, V)}{N}$$

where $X_j$ is the variance for respondent $j$.

Given our distributional assumptions, this operation is equivalent to maximizing expression (21).

$$I \frac{rX_j(N/V)}{J} \sum_{j=1}^{J} rX_j(N/V)$$

where $I$ is the number of respondents for which $n^* > M$. Expression (21) is maximized when $I = J$, yielding the estimate for $V$ given by expression (22).

$$V = \max \{ \frac{\sigma^2}{J} \}$$

So, we must maximize expression (23).

$$I \frac{rX_j(N/V)}{J} \sum_{j=1}^{J} rX_j(N/V)$$

Taking logs, we obtain expression (24).

$$\sum_{j=1}^{J} \log r - \frac{r(\sigma^2/M)}{N} X_j$$

The estimate for $r$, given by expression (25), follows.

$$r = J/N \sum_{j=1}^{J} X_j$$

References


Wald, A. (1947), Sequential Analysis (New York: John Wiley and Sons).