Marketing Analytical Framework

Eric Lin

上海 华东师范大学

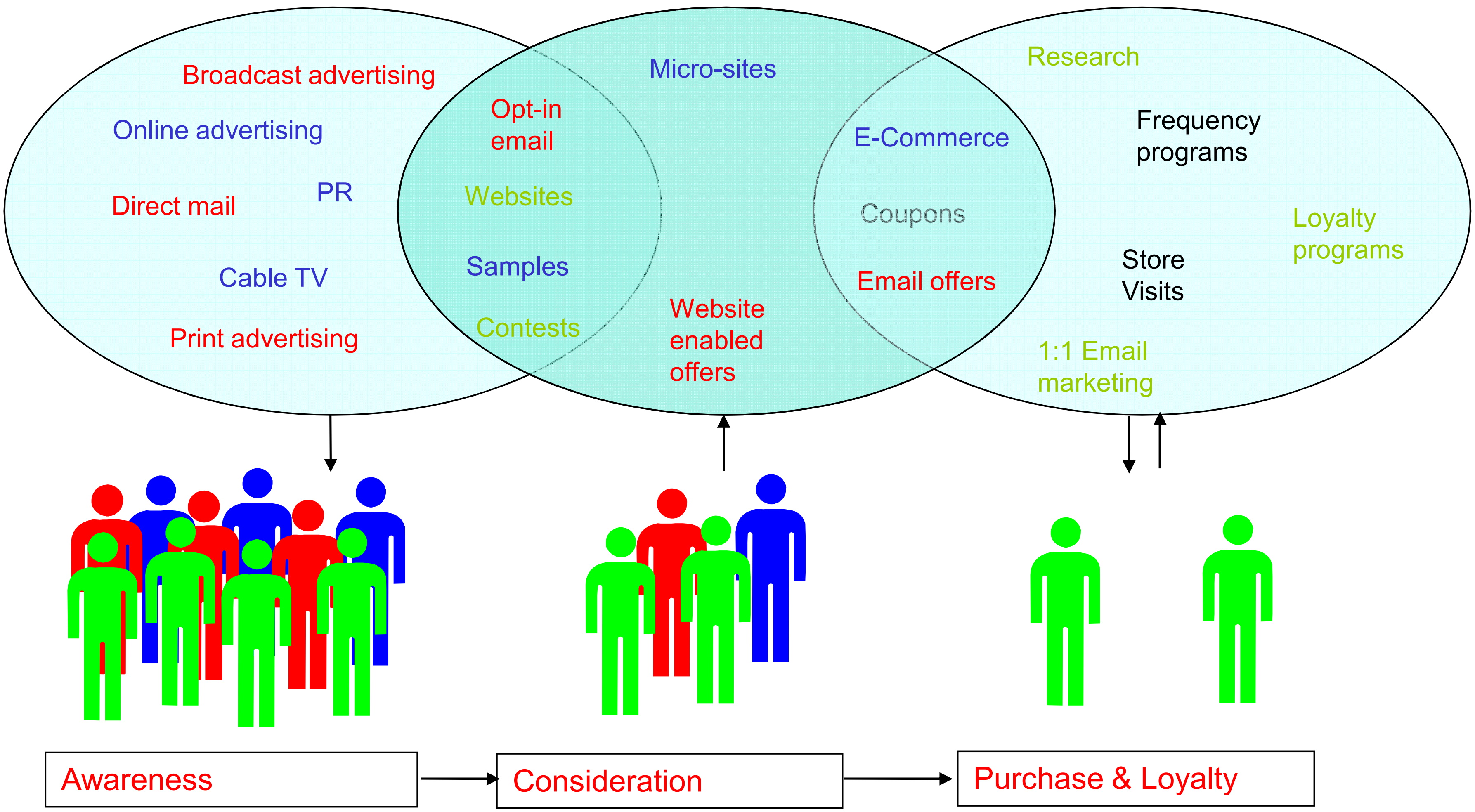
2009/12/13

Agenda

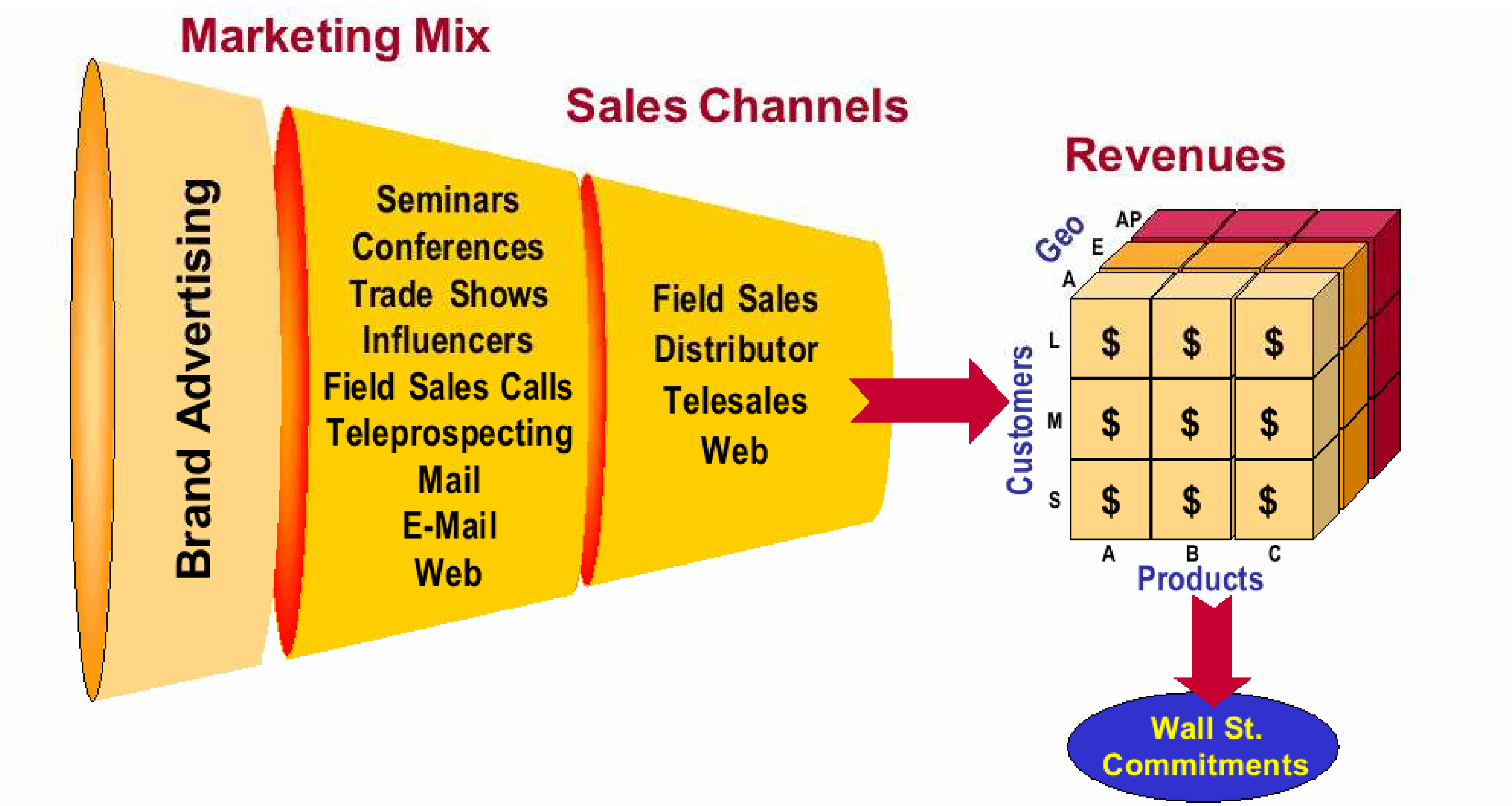
- Marketing Analytical Framework
- Direct Marketing Application
- Customer Base Analysis

Marketing Analytical Framework

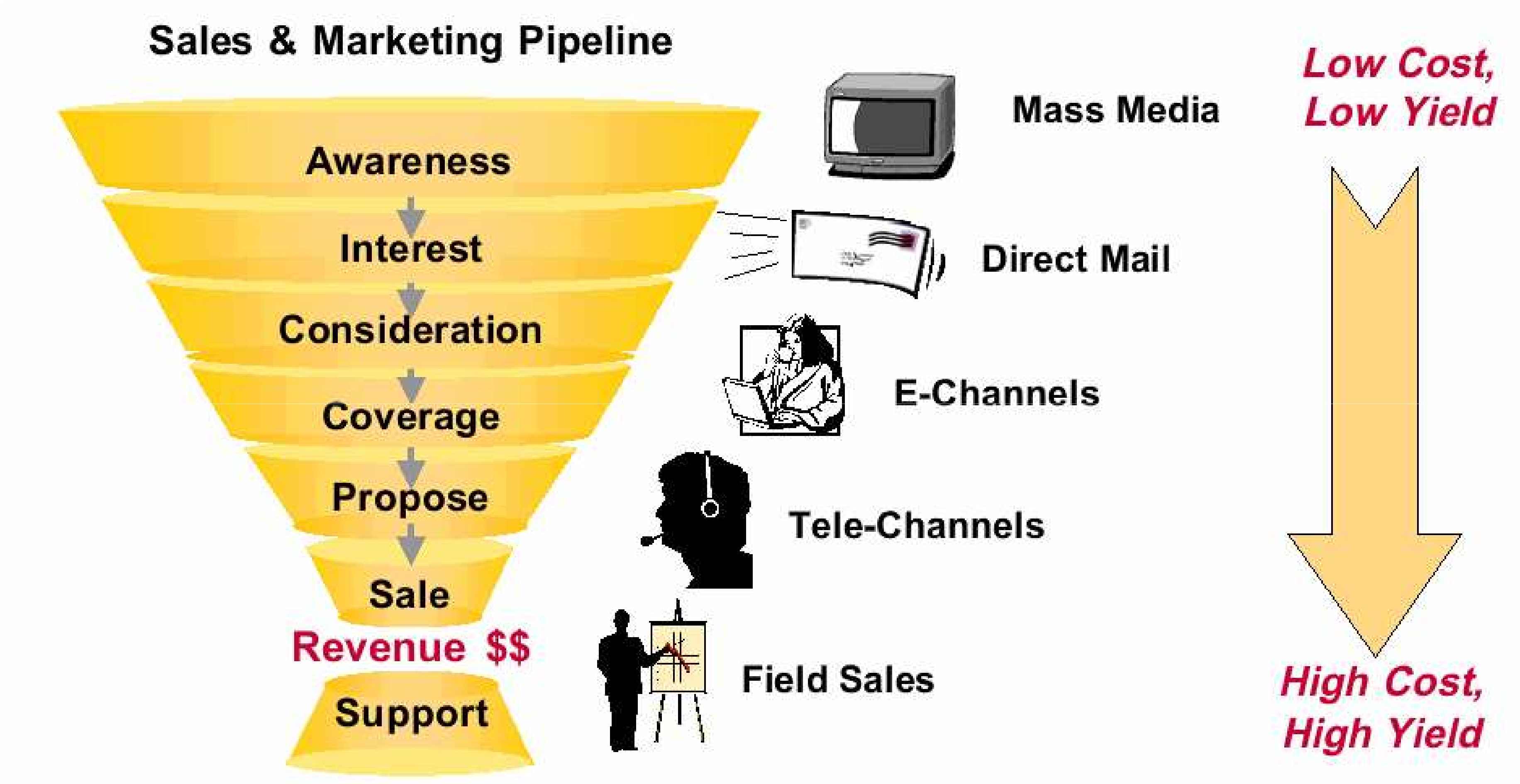
Marketing Today



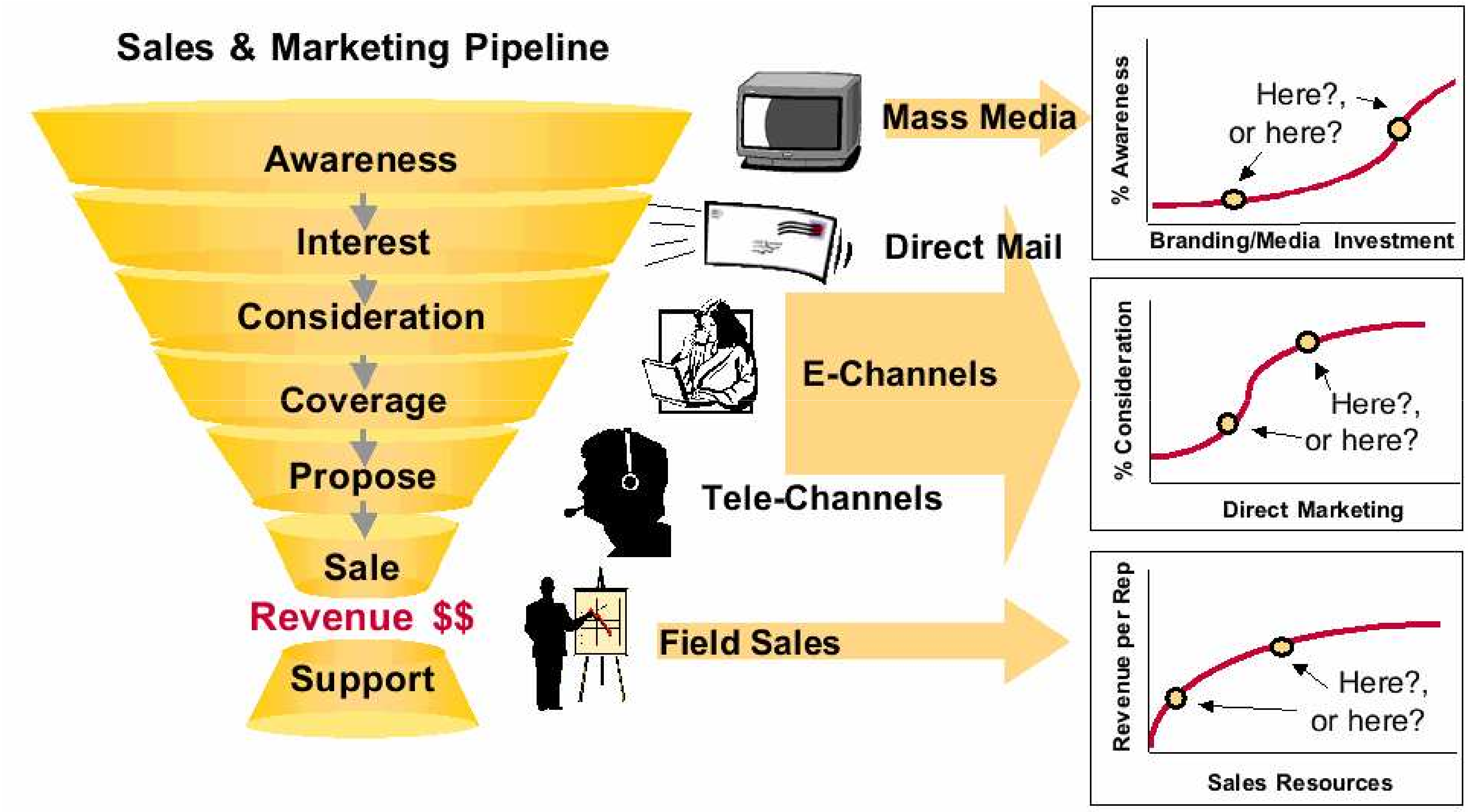
Management Challenge : Marketing Mix



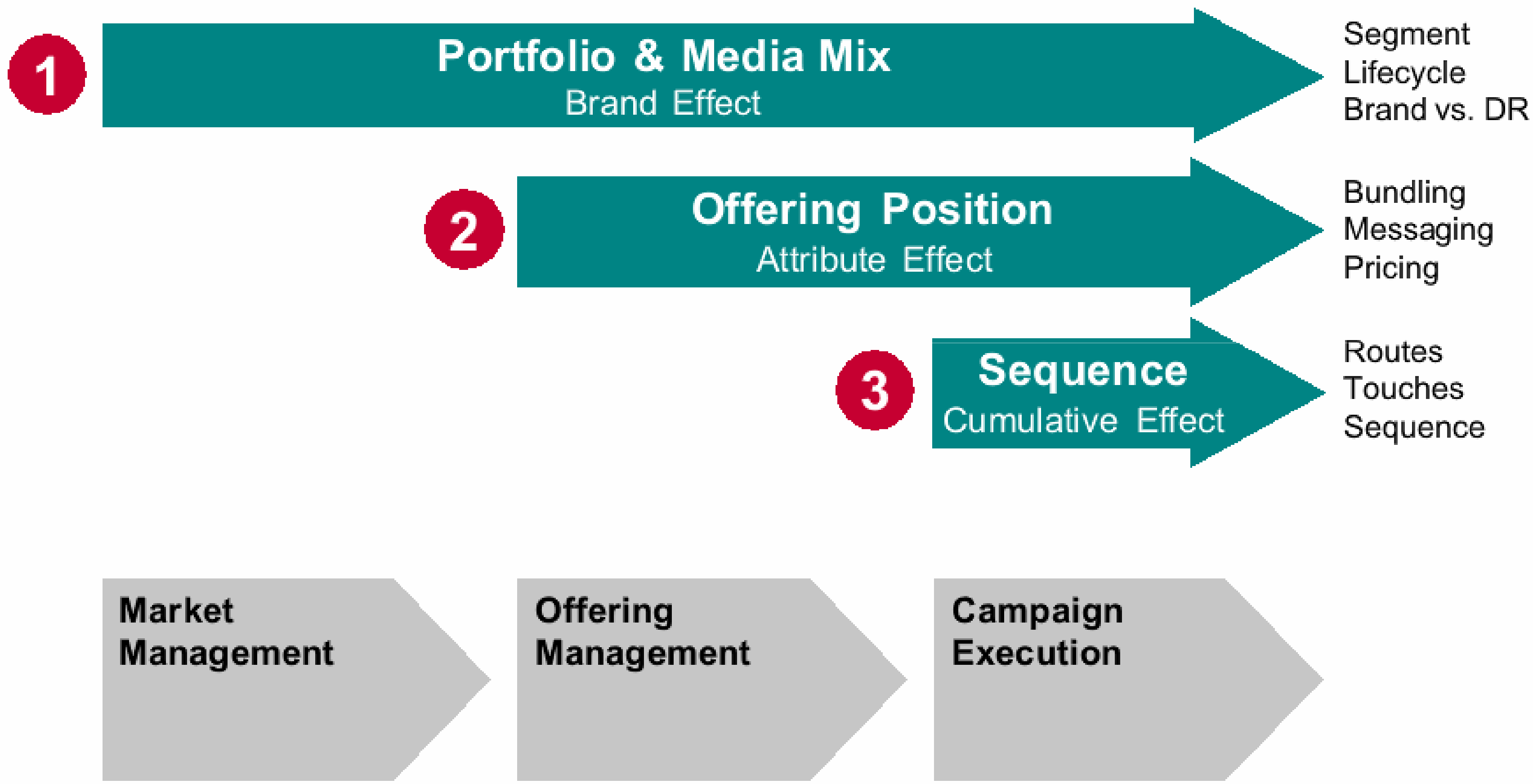
Mix & Sequence Varies Through The Pipline



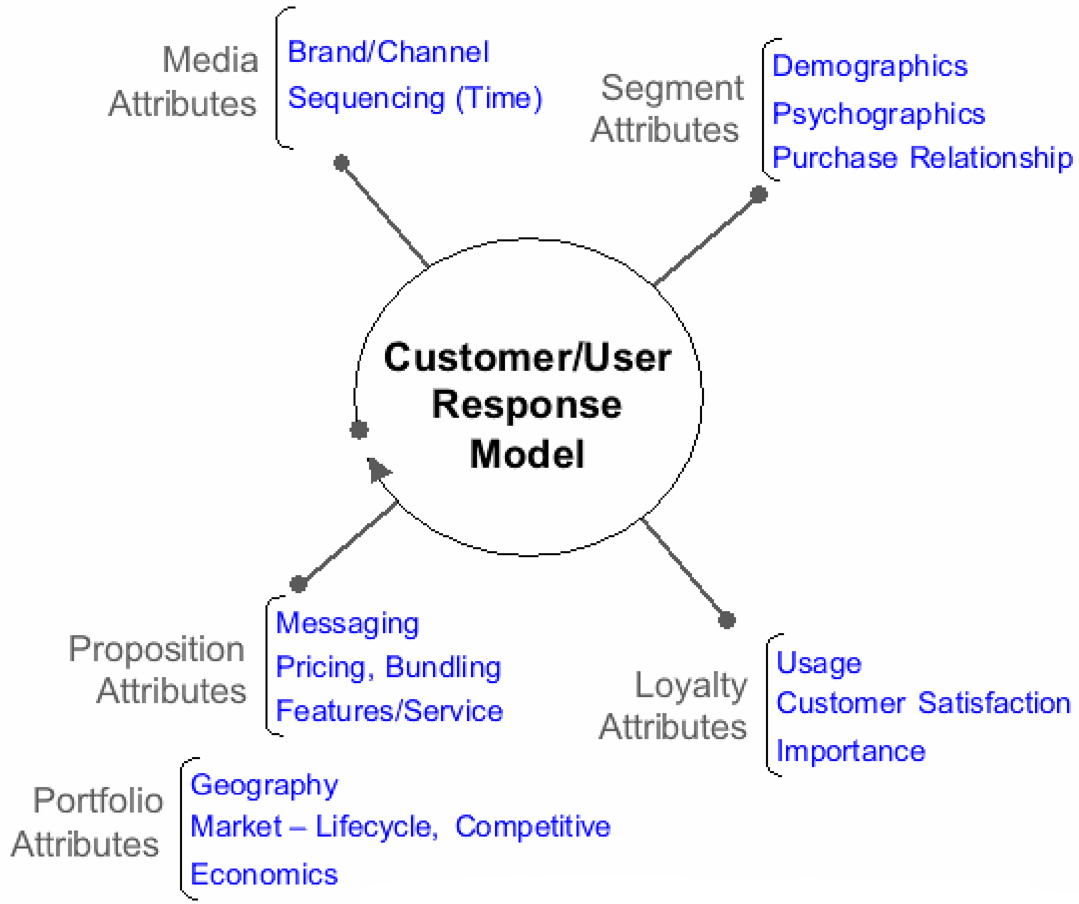
Marketing Optimization: Growing The Pipline



Interactive Effect Drive Optimization



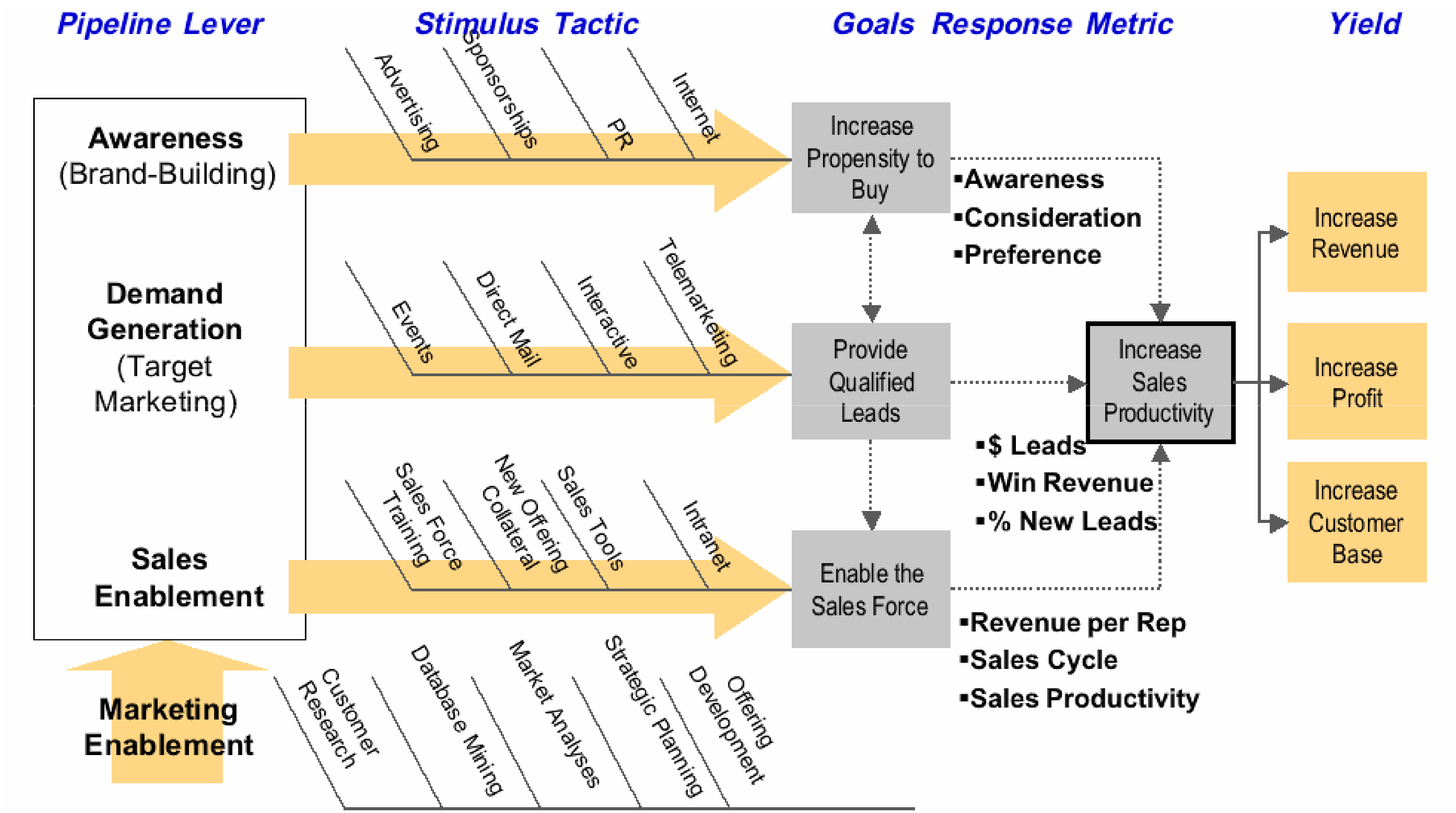
Marketing "Portfolio Management" Framework



What Are We Trying To Optimize?

- Awareness
- Consideration
- Purchase (Revenue)
- Profit
- Life-time value (retention)

Analytical Framework



Direct Marketing Application

The "Segmentation" Approach

1. Divide the customer list into a set of (homogeneous) segments.
2. Test customer response by mailing to a random sample of each segment.
3. Rollout to segments with a response rate (RR) above some cut-off point,

$$\text{e.g., RR} > \frac{\text{cost of each mailing}}{\text{unit margin}}$$

- A consumer durable product (unit margin = $161.50, mailing cost per 10,000 = $3343)
- 126 segments formed from customer database on the basis of past purchase history information
- Test mailing to 3.24% of database

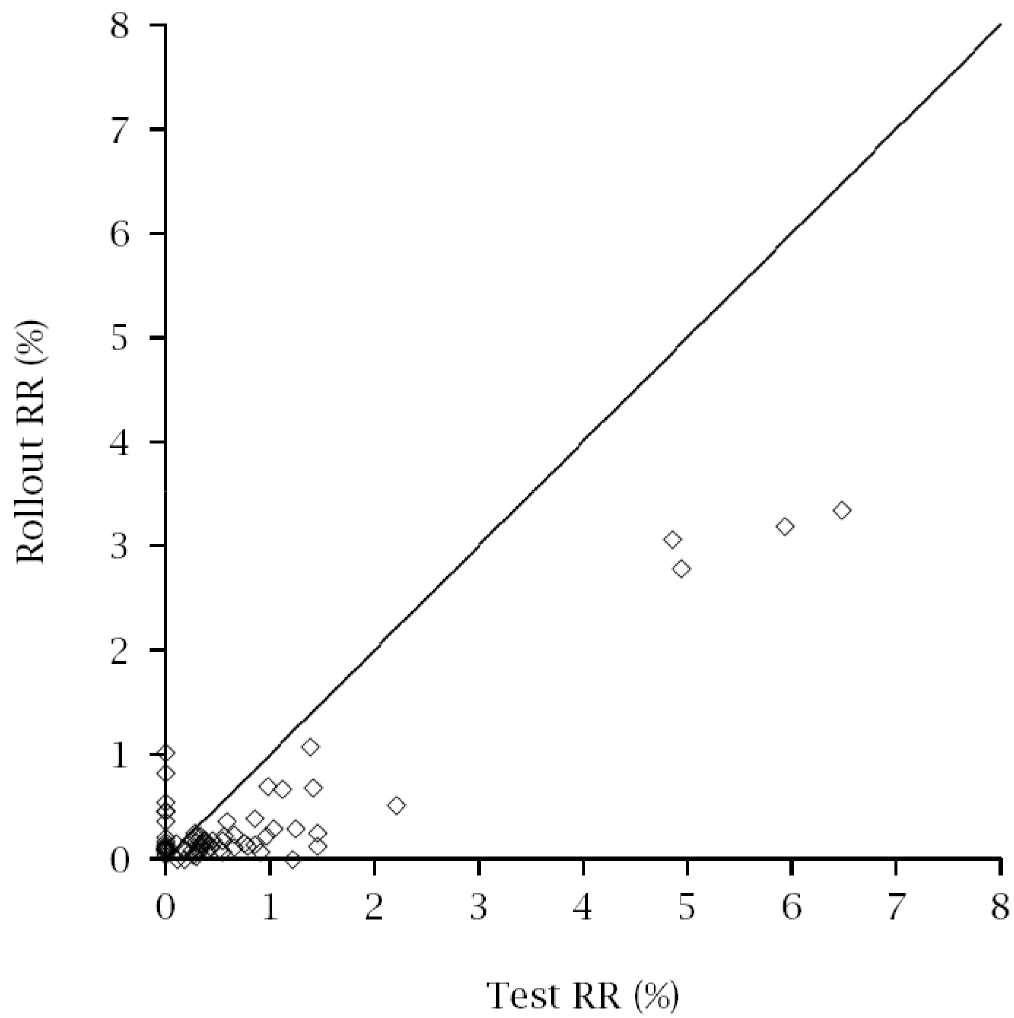
Standard approach:

- Rollout to all segments with

$$\text{Test RR} > \frac{3343/10,000}{161.50} = 0.00207$$

- 51 segments pass this hurdle

Test vs. Actual Response Rate



Modeling Objective

Develop a model that leverages the whole data set to make better informed decisions.

Model Development

Notation:

$$N_s = \text{size of segment } s \ (s = 1, \ldots, S)$$
$$m_s = \text{\# members of segment } s \text{ tested}$$
$$X_s = \text{\# responses to test in segment } s$$

Assume: All members of segment s have the same (unknown) response probability $p_s \Rightarrow X_s$ is a binomial random variable

$$P(X_s = x_s | m_s, p_s) = \binom{m_s}{x_s} p_s^{x_s} (1 - p_s)^{m_s - x_s}$$

Distribution of Response Probabilities

- Heterogeneity in p_s is captured using a beta distribution:

$$g(p_s) = \frac{1}{B(\alpha, \beta)} p_s^{\alpha-1}(1 - p_s)^{\beta-1}$$

- The beta function, $B(\alpha, \beta)$, can be expressed as

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

- The mean of the beta distribution is given by

$$E(p_s) = \frac{\alpha}{\alpha + \beta}$$

The Beta Binomial Model

The aggregate distribution of responses to a mailing of size m_s is given by

$$P(X_s = x_s | m_s) = \int_0^1 P(X_s = x_s | m_s, p_s)\, g(p_s)\, dp_s$$
$$= \binom{m_s}{x_s} \frac{B(\alpha + x_s, \beta + m_s - x_s)}{B(\alpha, \beta)}$$

Estimating Model Parameters

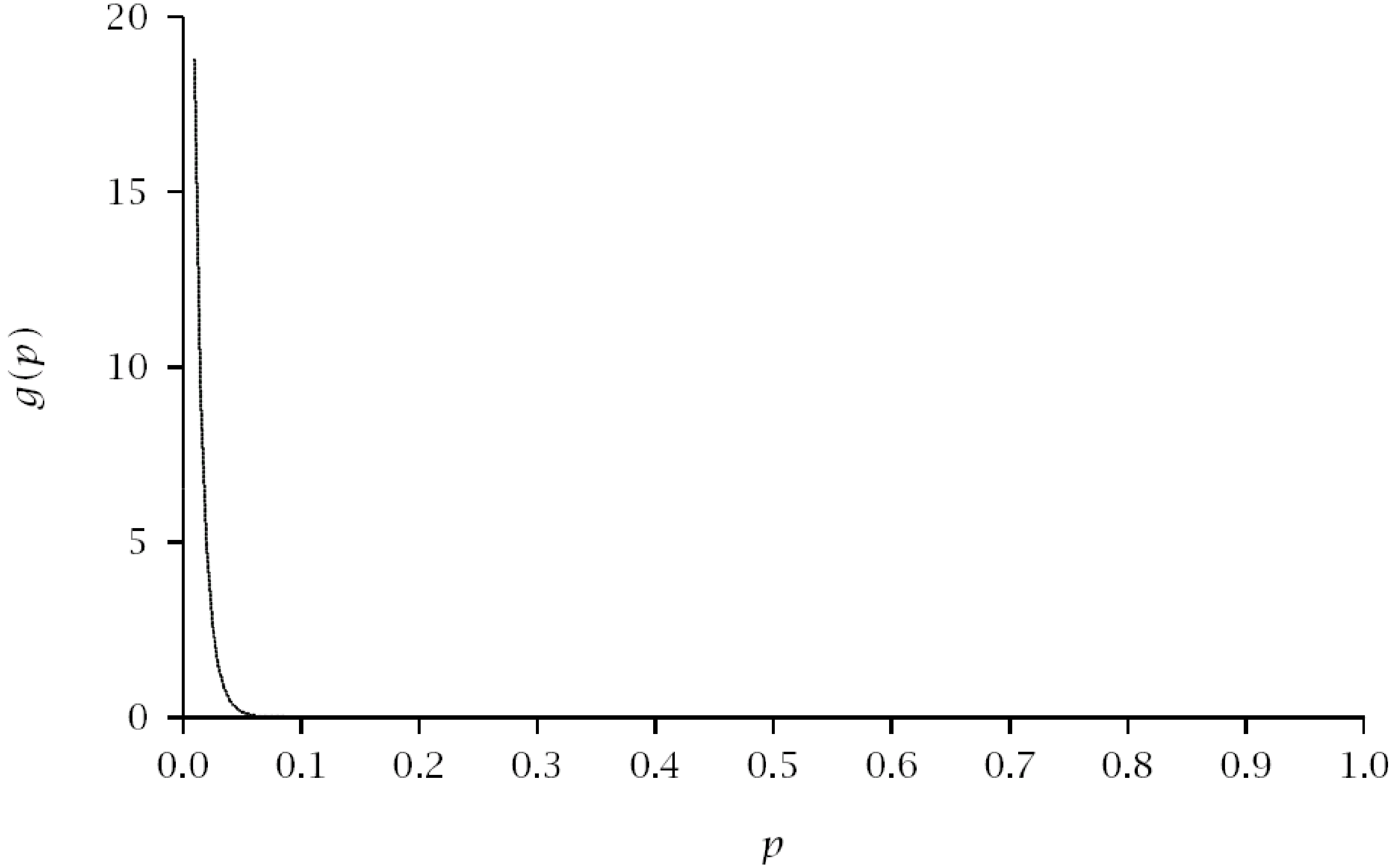
The log-likelihood function is defined as:

$$LL(\alpha, \beta | \text{data}) = \sum_{s=1}^{126} \ln[P(X_s = x_s | m_s)]$$

$$= \sum_{s=1}^{126} \ln\Bigg[\frac{m_s!}{(m_s - x_s)!\, x_s!} \underbrace{\frac{\Gamma(\alpha + x_s)\Gamma(\beta + m_s - x_s)}{\Gamma(\alpha + \beta + m_s)}}_{B(\alpha+x_s, \beta+m_s-x_s)} \underbrace{\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}}_{1/B(\alpha,\beta)} \Bigg]$$

The maximum value of the log-likelihood function is $LL = -200.5$, which occurs at $\hat{\alpha} = 0.439$ and $\hat{\beta} = 95.411$.

Estimated Distribution of p



$$\hat{\alpha} = 0.439, \hat{\beta} = 95.411, \bar{p} = 0.0046$$

Applying the Model

What is our best guess of p_s given a response of x_s to a test mailing of size m_s?

Intuitively, we would expect

$$E(p_s|x_s, m_s) \approx \omega \frac{\alpha}{\alpha + \beta} + (1 - \omega) \frac{x_s}{m_s}$$

Bayes Theorem

- The *prior distribution* $g(p)$ captures the possible values p can take on, prior to collecting any information about the specific individual.
- The *posterior distribution* $g(p|x)$ is the conditional distribution of p, given the observed data x. It represents our updated opinion about the possible values p can take on, now that we have some information x about the specific individual.
- According to Bayes theorem:

$$g(p|x) = \frac{f(x|p)g(p)}{\int f(x|p)g(p)\,dp}$$

Bayes Theorem

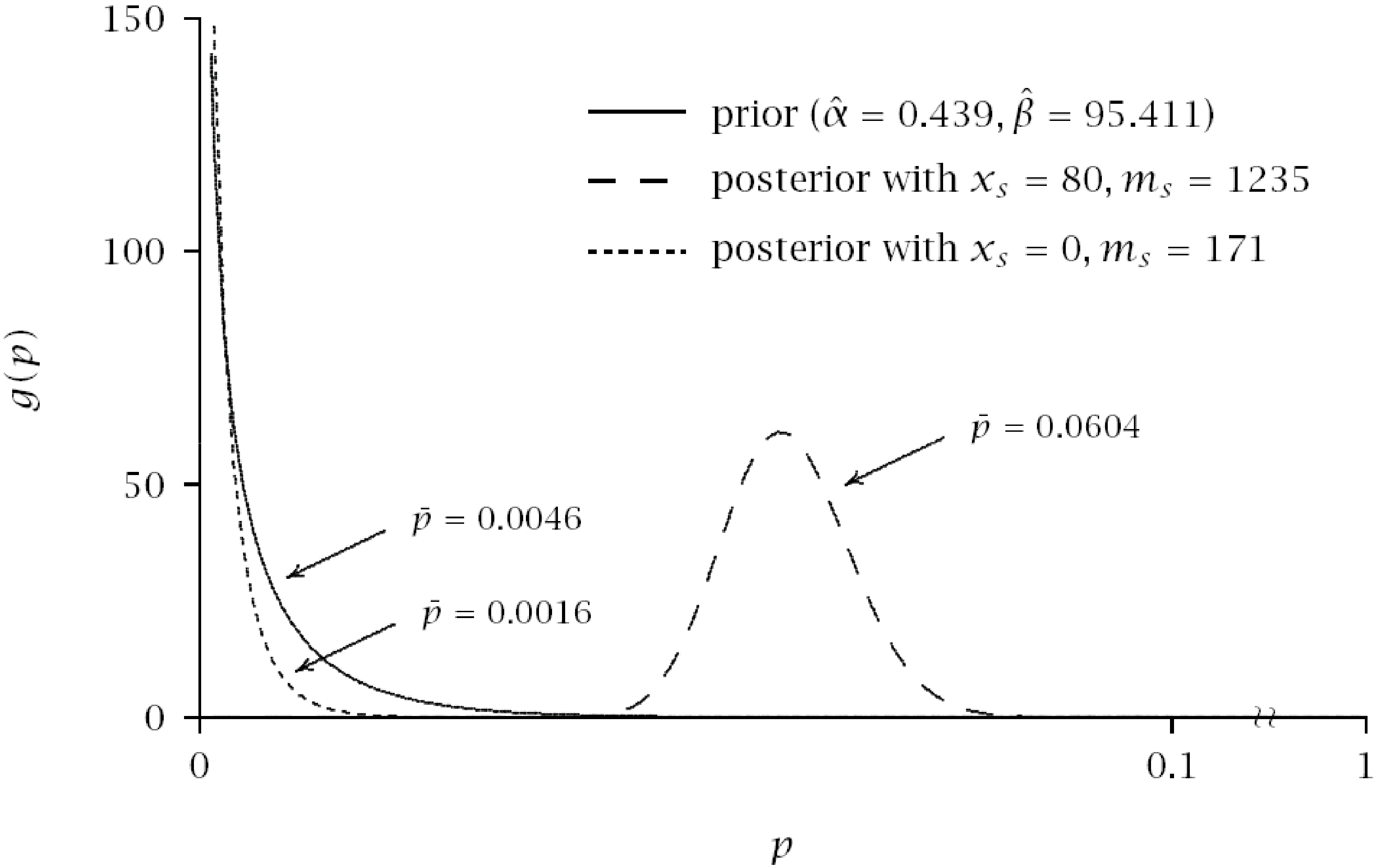
For the beta-binomial model, we have:

$$g(p_s|X_s = x_s, m_s) = \frac{\overbrace{P(X_s = x_s|m_s, p_s)}^{\text{binomial}}\,\overbrace{g(p_s)}^{\text{beta}}}{\underbrace{\int_0^1 P(X_s = x_s|m_s, p_s)\, g(p_s)\, dp_s}_{\text{beta-binomial}}}$$

$$= \frac{1}{B(\alpha + x_s, \beta + m_s - x_s)} p_s^{\alpha + x_s - 1}(1 - p_s)^{\beta + m_s - x_s - 1}$$

which is a beta distribution with parameters $\alpha + x_s$ and $\beta + m_s - x_s$.

Distribution of *p*



Applying the Model

Recall that the mean of the beta distribution is $\alpha/(\alpha+\beta)$. Therefore

$$E(p_s|X_s = x_s, m_s) = \frac{\alpha + x_s}{\alpha + \beta + m_s}$$

which can be written as

$$\left(\frac{\alpha+\beta}{\alpha+\beta+m_s}\right)\frac{\alpha}{\alpha+\beta} + \left(\frac{m_s}{\alpha+\beta+m_s}\right)\frac{x_s}{m_s}$$

- a weighted average of the test RR (x_s/m_s) and the population mean ($\alpha/(\alpha+\beta)$).
- "Regressing the test RR to the mean"

Model-Based Decision Rule

- Rollout to segments with:

$$E(p_s | X_s = x_s, m_s) > \frac{3343/10,000}{161.5} = 0.00207$$

- 66 segments pass this hurdle
- To test this model, we compare model predictions with managers' actions. (We also examine the performance of the "standard" approach.)

Customer Base Analysis

The simple models for three behavioral processes

- Timing → "when"
- Counting → "how many"
- "Choice" → "whether/which"

1. Each of these simple models has multiple applications
2. More complex behavioral phenomena can be captured by combining models from each of these processes

Further Applications: Timing Models

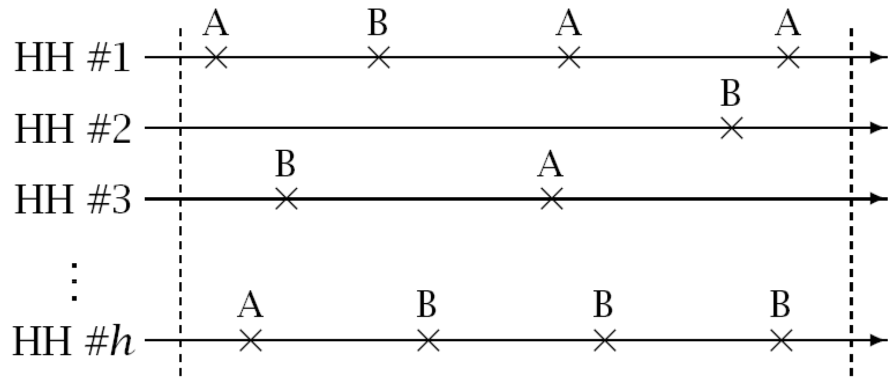
- Repeat purchasing of new products
- Response times:
  - Coupon redemptions
  - Survey response
  - Direct mail (response, returns, repeat sales)
- Customer retention/attrition
- Other durations:
  - Salesforce job tenure
  - Length of web site browsing session

Further Applications: Count Models

- Repeat purchasing
- Customer concentration (“80/20” rules)
- Salesforce productivity/allocation
- Number of page views during a web site browsing session

Further Applications: "Choice" Models

- Brand choice



- Media exposure
- Multibrand choice
- Taste tests (discrimination tests)
- "Click-through" behavior

Thanks and Any question?