STUDY DESIGNS

IN BIOMEDICAL RESEARCH

Design & Modeling Issues in

DEMAND CURVE ANALYSIS

The Family Smoking Prevention and Tobacco Control Act is a federal statute which was signed into law by President Obama on June 22, 2009. The Act gives the Food and Drug Administration (FDA) the power to regulate the tobacco industry.

After the Family Smoking Prevention and Tobacco Control Act passed, tobacco research have been going even stronger and branching into more directions. One of the major areas is Product Liability – part of a new territory called “Regulatory Science”. As part of those efforts, many focus on the Modeling and Data Analysis of the Demand Curve in recent years.

THE DEMAND CURVE

A fundamental concept of consumer demand, in Behavioral Economics, is the Demand Curve relating the consumption of a commodity (C, dependent variable) to its price (P, the independent variable). According to the theory, the consumption of most goods will decrease with increases in price (Watson and Holman, 1977).

ELASTICITY

At the discrete level, a section of the demand curve is characterized by a parameter called Elasticity (E) which is defined as the ratio of two rates/proportions (% over %):

$$E = \frac{\dfrac{(C_2 - C_1)}{(1/2)(C_1 + C_2)}}{\dfrac{(P_2 - P_1)}{(1/2)(P_1 + P_2)}}$$

“Elasticity” could be used to compare liability between products; e.g. with the same price increase, consumption of one tobacco product would reduce faster than that of the other. For that example, the difference could represent different levels of dependency or addiction.

ELASTICITY on Continuous Scale

For a point on the demand curve, i.e. continuous scale, the elasticity E becomes:

$$E = \frac{\dfrac{(C_2 - C_1)}{(1/2)(C_1 + C_2)}}{\dfrac{(P_2 - P_1)}{(1/2)(P_1 + P_2)}} \longrightarrow E = \left[\frac{P}{C}\right]\left[\frac{dC}{dP}\right] = \frac{d[\ln C]}{d[\ln P]}$$

which represents the slope on the demand curve when both price (P) & consumption (C) are expressed on the log scale (we might but do not have to graph the curve with axes marked on log scale).

DEMAND CURVE FOR TOBACCO RESEARCH

The demand curve established for food consumption has been adopted for use in tobacco research in areas of product liability and relative reinforcing efficacy (RRE), a concept in psychopharmacological research (Bickel and Madden 1999). It has also been used in studies of drugs, like cocaine. There are studies both in humans (surveys of smokers) & animals (experiments with rats).

AN ANIMAL EXPERIMENT:

Human research suggests that there are sex differences in the addiction-related behavioral effects of nicotine; a study was conducted to examine this issue in rats:

- ❖ **Male and female rats were trained to self-administer nicotine (0.06 mg/kg) under a FR 3 schedule during daily 23-hour sessions.**
- ❖ **Rats were then exposed to saline extinction and reacquisition of NSA, followed by weekly reductions in the unit dose (0.03 to 0.00025 mg/kg) until extinction levels of responding were achieved.**
- ❖ **Fifteen rats (8 males, 7 females) were tested at 8 doses: 0.03, 0.02, 0.01, 0.007, 0.004, 0.002, 0.001, 0.0005, mg/kg/infusion.**

A SURVEY OF SMOKERS

- **Data are collected by the cigarette purchase task (CPT) survey, also called TPT, in which participants were asked to respond to the following set of questions**
- **How many cigarettes would you smoke if they were______ each?: 0¢ (free), 1¢, 5¢, 13¢, 25¢, 50¢, $1, $2, $3, $4, $5, $6, $11, $35, $70, $140, $280, $560, $1,120.**
- **This set of questions are asked during an online survey in the preceding order until the respondents gives "0" as an answer, then no more further questions will be asked.**

HURSH'S FIRST MODEL

Collecting data on the consumption of foods, Hursh et al. (1989) found empirical evidence that demand elasticity is a linear function of price ($E = b - aP$) leading to a specific equation of the demand curve. This first curve used in the study of foods a very simple model, a straight line.

There were recent efforts by Hursh and Silberberg to form a one-parameter model (2008):

(1) They set "the goal of defining a single parameter for indexing the rate of change in elasticity of demand" because the linear elasticity model fails "to define essential value";

(2) They followed the observation by Allen (1962) that a demand curve is "downward sloping in price-consumption space"; then chose one of the eight possible equations provided by Allen. These are individual curves, not a population or global curve.

HURSH-SYLBERBERG MODEL

Hursh-Sylberberg model is expressed as:

$$\ln C = \ln C_{(0)} + k[\exp(-\alpha P) - 1]$$

$$\ln C_{(0)} = \ln \underset{P \to 0}{C}$$

THE HURSH-SILBERBERG MODEL:

P is Price
Q is Demand/Consumption
$C_{(0)}$ is Level of demand when price approaches 0
k is related to the range of Q
α is a measure of elasticity
Note: We use two different notations: C_0 is the current consumption and $C_{(0)}$ denotes consumption at price zero – as in the model.

From the first model (1998) to the second model (2008), the focus shifts from the shape of "elasticity curve" (a line) to the "demand curve". And the resulting model has become the current standard of the field used in numerous publications in several products.

ESTIMATION OF PARAMETERS

By treating the Hursh-Silberberg model as a non-linear regression model and supplementing with some distribution for the error term, we can estimate α, k, and $C_{(0)}$ and obtain their standard errors. Computation can be implemented using computer programs (Prism, SAS). Estimates may be different because of different estimation techniques but differences are small (SAS is standard for statisticians but Prism is very popular with investigators in applied fields.

Issue: DATA QUALITY

If data from certain subject do not fit well (low R^2), some investigators would exclude the subject. But this is a rather tricky step: How low is low? How to defend for setting certain specific threshold? .5? .8? It is more problematic, especially with human data.

For CPT, questions start at zero; responses to questions at prices below a smoker's current price may be shaky because some smokers might feel guilty about their habits and not provide consumption above the current consumption. For these subjects, their curves start with a "flat" section, only go down after the current price; (left) truncating the first part would improve the fit . This suggests a "prospective design" and the need to standardize the Demand Curve.

“PROSPECTIVE” DESIGN

If our interests are in elasticity parameters, why do we want to know C_0 ?

(1) In animal studies, after training the rats for self administration, experiment starts at a price P_0 – raising to prices which are multiple of P_0 ;

(2) CPT survey could start with the question “how much are you paying” to obtain current price P_0; then the online survey could be programmed to follow with prices which are multiple of price P_0 just obtained.

STANDARDIZED DEMAND CURVE

The Demand Curve relates the consumption (C, dependent variable – on vertical axis) to its price (P, the independent variable – on the horizontal axis).

For both animal and human data, the current consumption level C_0 for each subject is available, we can easily form a Standardized Demand Curve with C/C_0 as the dependent variable – on vertical axis.

STANDARDIZED DEMAND CURVE AS A SURVIVAL CURVE

In the Standardized Demand Curve, if we denote

$$t = \ln(P/P_0)$$

$$S(t) = \frac{C}{C_0}$$

Each individual could be viewed as a "Survival Curve" with "survival rate" $S(t) = C/C_0$ going down from 1.0 as "time" t increases. This same curve serves as a global representation of "consumption reduction" versus "price increase"

WHAT IS ELASTICITY?

$$S(t) = \frac{C}{C_0} \;\&\; t = \ln(P/P_0) = \ln P - \ln P_0$$

$$\ln[S(t)] = \ln(C) - \ln(C_0)$$

$$h(t) = -\frac{d}{dt}\ln[S(t)] = -\frac{d(\ln C)}{d(\ln P)}$$

$$= -\text{Elasticity}$$

If we view the Standardized Demand Curve as a survival curve, Elasticity Function is simply the negative of the Hazard Function.

What motivated the import of ideas from behavioral economics is the concept Elasticity or Elasticity Function which is simply the negative of the Hazard Function if the Standardized Demand Curve is viewed as a Survival Curve. However, what else do we have besides Hursh-Sylberberg model ? Or, if Hursh-Sylberberg model could be viewed as a survival curve?

The finding that we could view the Standardized Demand Curve as a Survival Curve (and the Elasticity Function could be derived from the Hazard Function), would open up possibilities for modeling and more efficient strategies for data analysis.

STRATEGY

1) Aiming at the Elasticity Function
2) Modeling the Standardized Demand Curve – as a Survival Curve;
3) Estimating its parameters;
4) Using estimated parameters to form the Elasticity Function, from the Hazard Function, as the ultimate products.

Possible Model #1: WEIBULL

$$\text{Standardized Demand Curve} \quad S(t) = \exp[-(\alpha t)^{\beta}]$$

$$\text{Elasticity} \quad E(t) = -\alpha\beta(\alpha t)^{\beta-1}$$

Could it be more simple? Yes, if data fits the Exponential (β=1), the standardized demand curve depends on one constant.

Possible Model #2: LOG-LOGISTIC

$$\text{Standardized Demand Curve} \quad S(t) = \frac{1}{1+(\alpha t)^{\beta}}$$

$$\text{Elasticity} \quad E(t) = -\frac{\alpha^{\beta}\beta t^{\beta-1}}{1+(\alpha t)^{\beta}}$$

Another simple two-parameter model; the question is how to measure goodness-of-fit and choose the better model.

DATA ANALYSIS STRATEGIES

- Two choices:
- Starting with individual curves, then combing results to form population curve.
- Going right to population curve, and treat individual data as repeated observations.
- We'll take and illustrate the first approach which is much more simple to see and to implement – using Simple Linear Regression.

Model #1: WEIBULL

$$t = \ln(P/P_0)$$

$$S(t) = \frac{C}{C_0}$$

$$S(t) = \exp[-(\alpha t)^{\beta}]$$

$$E(t) = -\alpha\beta(\alpha t)^{\beta-1}$$

→

$$\ln[-\ln S(t)] = \beta\ln\alpha + \beta\ln t$$

$$\ln[-\ln \frac{C}{C_0}] = \beta\ln\alpha + \beta\ln[\ln(P/P_0)]$$

We have a simple linear regression after two double log transformations . We combine individual results by calculating weighted averages of slopes and intercepts using inverse of variance as the weight; and use these weighted averages to form Standardized Demand & Elasticity functions.

Model #2: LOG-LOGISTIC

$$t = P$$

$$S(t) = \frac{C}{C_0}$$

$$S(t) = \frac{1}{1 + (\alpha t)^{\beta}}$$

$$E(t) = -\frac{\alpha^{\beta}\beta t^{\beta-1}}{1 + (\alpha t)^{\beta}}$$

$\longrightarrow$

$$\ln\left(\frac{1 - S(t)}{S(t)}\right) = \beta\ln\alpha + \beta\ln t$$

$$\ln\left[\frac{1 - \frac{C}{C_0}}{\frac{C}{C_0}}\right] = \beta\ln[\ln(P/P_0)]$$

Again, we have a simple linear regression after a logit and a double log transformations. We combine individual results by calculating weighted averages of slopes and intercepts using inverse of the variance as the weight; and use these weighted averages to form Standardized Demand & Elasticity functions.

For each subject, and assuming each model, we have a Simple Linear Regression (after different data transformations). The goodness-of-fit of the line is measure by the conventional R^2 (Coefficient of Determination); we could average them out, across subjects, to obtain an Overall R^2. Then use this Overall R^2 to judge goodness-of-fit of each model and select as the better model, the one with larger Overall R^2. For example:

$$R^2 = \frac{SSR}{SST}$$

$$\text{Overall } R^2 = \frac{\sum SSR}{\sum SST}$$

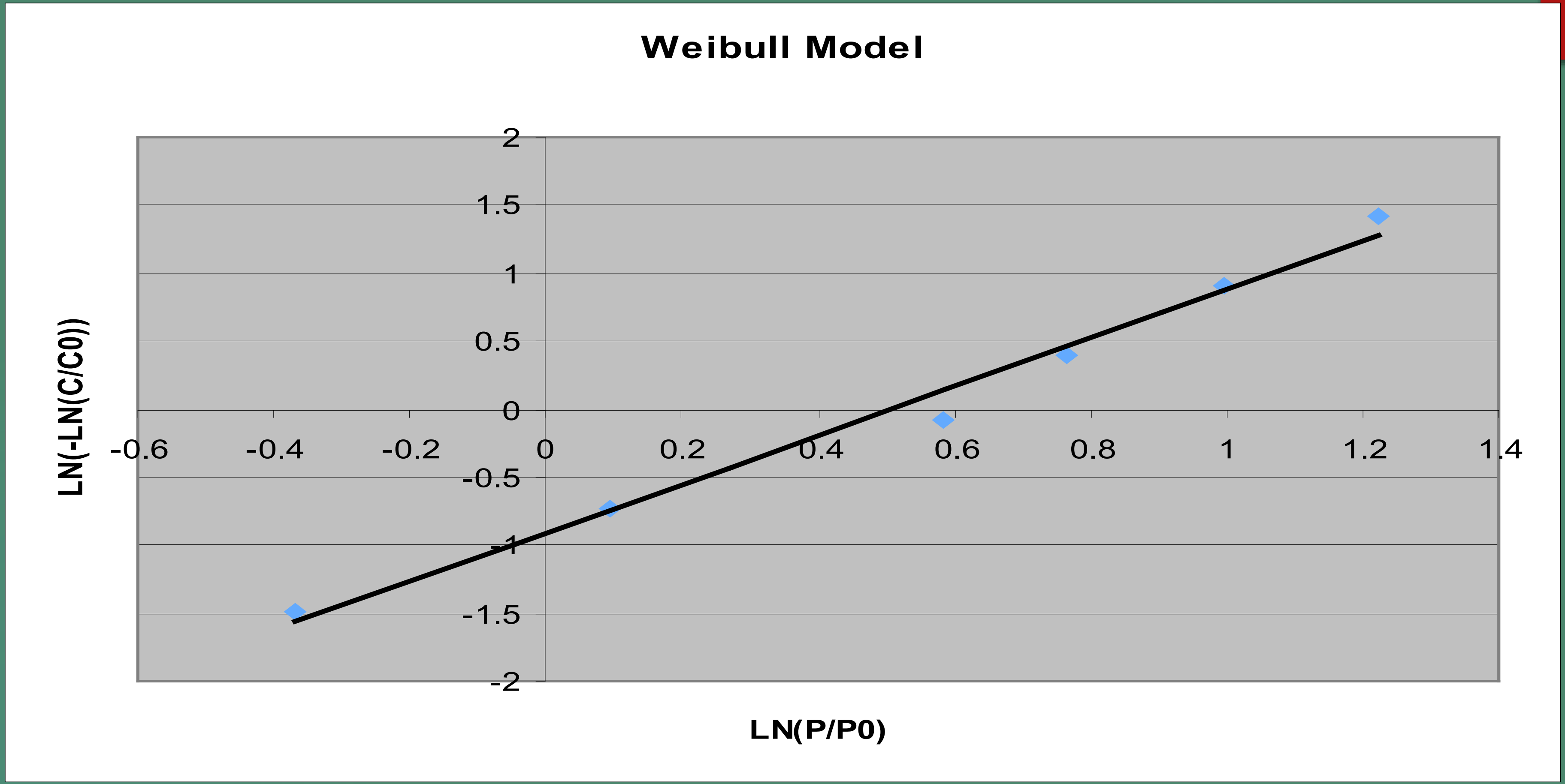
A TYPICAL ANIMAL EXPERIMENT

- Animal and human research suggests that there are sex differences in the addiction-related behavioral effects of nicotine.
- A study was conducted to examine this issue in rats. A nicotine reduction policy is modeled by arranging progressive decreases in the unit dose of nicotine available for self-administration, similar to human studies that have examined progressive reduction of cigarette nicotine content. Fifteen rats included: 8 males, 7 females
- Doses are: 0.03, 0.02, 0.01, 0.007, 0.004, 0.002, 0.001, 0.0005, mg/kg/infusion.

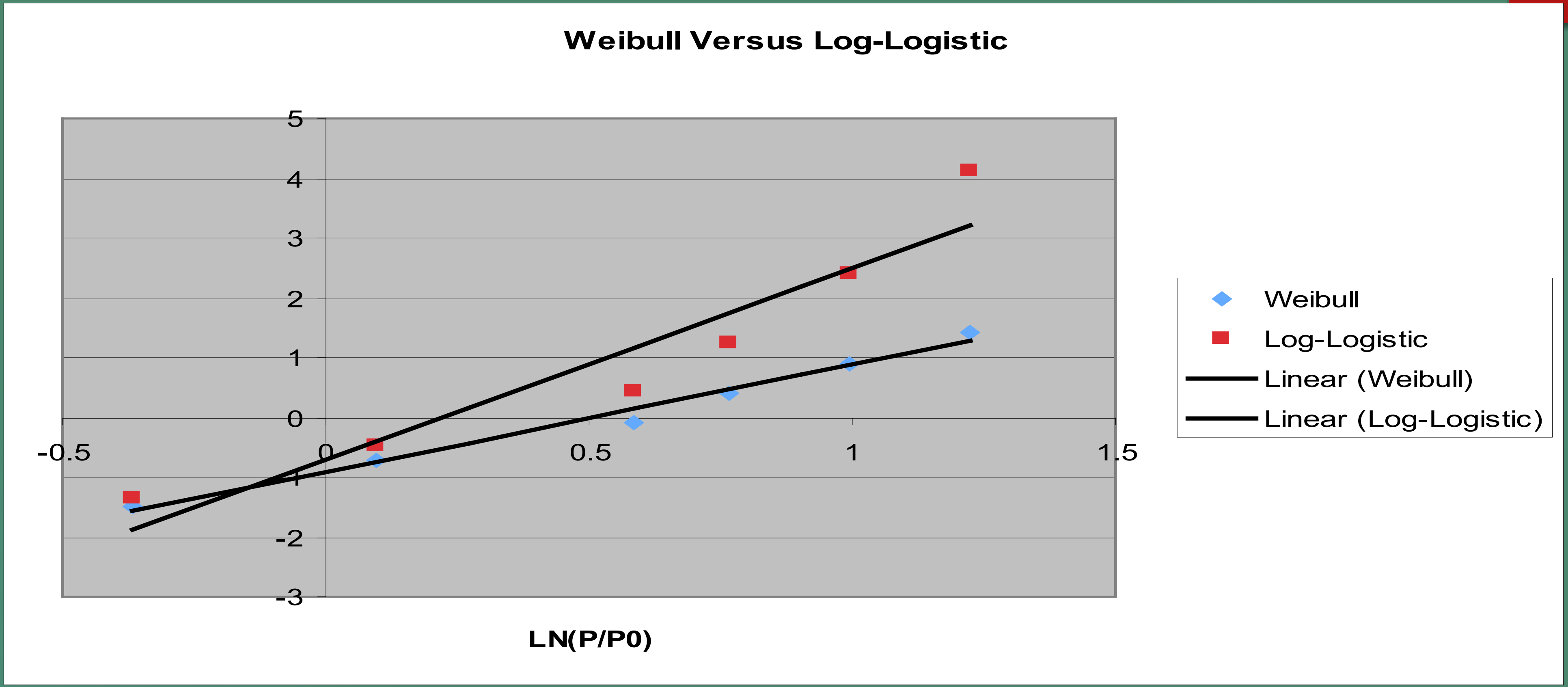
NUMERICAL EXAMPLE: RATS DATA

Price	Males							
50	2.0220	1.4820	1.4400	2.2200	1.3800	1.6800	2.0820	1.6560
100	1.6110	1.0800	0.9810	1.4610	1.0110	1.3200	1.7490	0.7710
150	1.2460	0.8200	0.8140	1.0400	0.9000	1.0460	1.3940	0.7200
300	0.8000	0.6130	0.4600	0.6430	0.5870	0.6000	0.6170	0.4600
429	0.4571	0.1589	0.1449	0.4179	0.4809	0.4760	0.1281	0.2751
750	0.1692	0.0588	0.0520	0.2200	0.2200	0.2492	0.0172	0.0960
1500	0.0320	0.0106	0.0140	0.0346	0.0880	0.0834		0.0180
3000				0.0117	0.0290			
6000					0.0057			
12000					0.0022			

Price	Females						
50	3.0000	2.2200	1.6200	3.3600	1.9980	1.8420	2.3220
100	1.4400	0.9690	0.8310	1.8990	1.1700	1.2210	1.3500
150	1.0260	0.9000	0.7340	1.1260	1.0000	1.0940	0.9260
300	0.6830	0.2630	0.2870	0.6830	0.5500	0.7200	0.6700
429	0.1680	0.1470	0.0959	0.5670	0.4571	0.5159	0.4501
750	0.0652	0.0320	0.0428	0.3348	0.2188	0.2732	0.2320
1500			0.0094	0.0586	0.0606	0.1314	0.0400
3000					0.0193	0.0227	0.0157
6000					0.0108	0.0089	
12000							



Weibull Model fits well



Weibull Model fits better than Log-Logistic Model

RESULTS

Males

Intercept	-0.911	-0.667	-0.544	-0.452	-0.891	-0.883	-0.995	-0.123
SE(Intercept)	0.08	0.158	0.122	0.081	0.102	0.034	0.125	0.151
Slope	1.793	1.752	1.651	1.416	1.56	1.558	2.463	1.115
SE(Slope)	0.104	0.205	0.158	0.091	0.093	0.044	0.194	0.197
R^2	0.987	0.948	0.965	0.98	0.976	0.997	0.982	0.889
Global Parameters								
Alpha = 0.603								
Beta = 1.585+/-0.032								
R^2 = 0.971								

Females

Intercept	0.015	0.066	-0.108	-0.139	-0.378	-0.662
SE(Intercept)	0.133	0.128	0.11	0.085	0.088	0.113
Slope	1.195	1.245	1.373	1.104	1.261	1.389
SE(Slope)	0.207	0.199	0.143	0.11	0.088	0.113
R^2	0.917	0.929	0.958	0.962	0.972	0.962

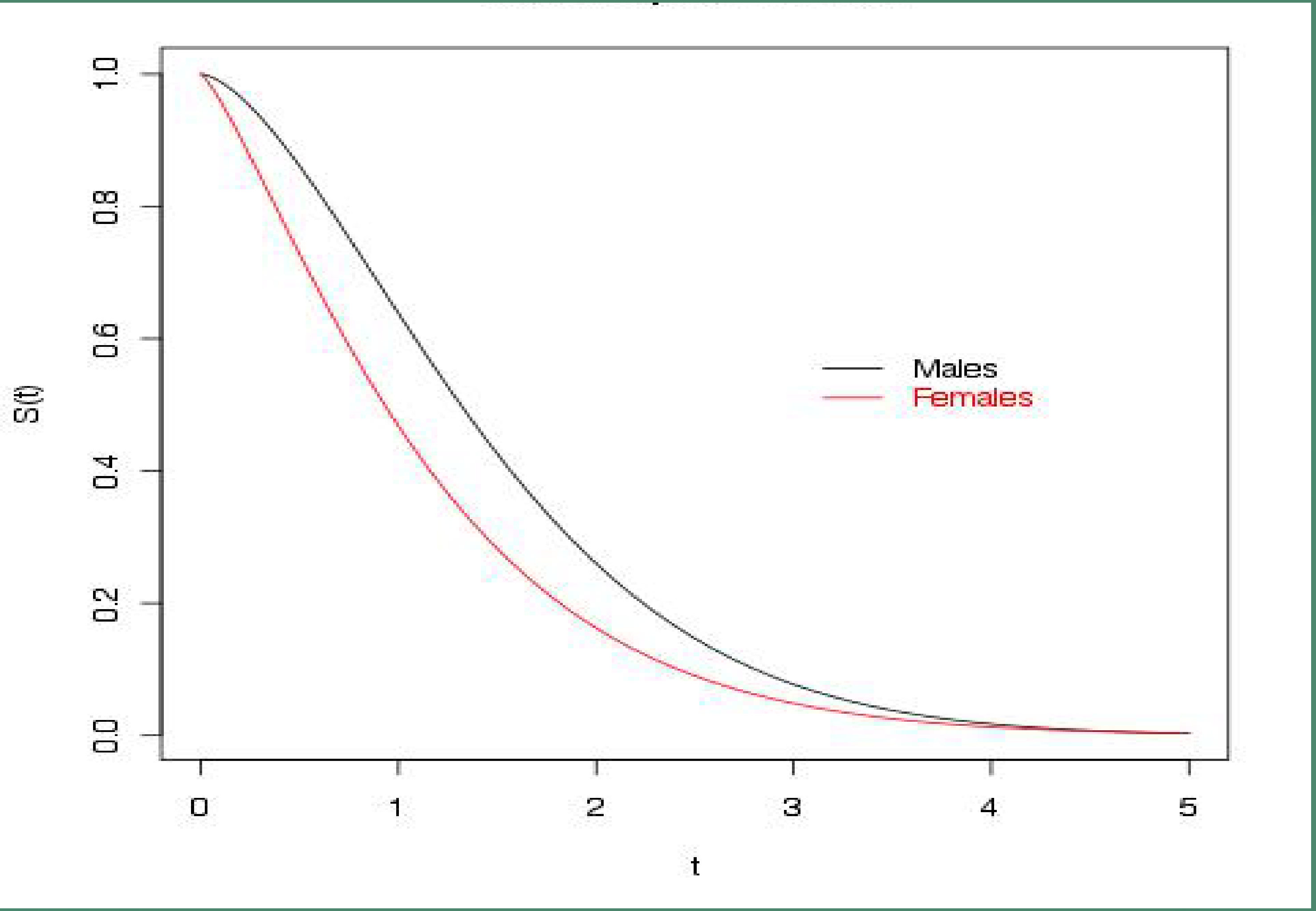
Global Parameters:

Alpha = 0.803

Beta = 1.258+/-0.047

R^2 = 0.956

STANDARDIZED DEMAND CURVES



(1) The shape of the curves is different from that shown on Hursh-Sylberberg second model; it's concave, not convex;

(2) The sloping down of these curves is not "exponential"; the shape parameter, β, is not equal to 1 (1.585 ± 0.032 for males and 1.248 ± 0.047 for females);

(3) The curve for female rats dropping down faster first at lower prices then the difference becomes smaller as price increases past 500-1000 units.

"HALF-BREAK POINT"

By setting ($C/C_0 = ½$), we would obtain the Price at 50% consumption reduction; let call it $P_{50\%}$:

$$P_{50\%} = P_0 \exp\{\frac{1}{\alpha}\exp[\frac{\ln(\ln 2)}{\beta}]\}$$

$$P_0 = 50 \text{ for males and females}$$

$$\text{Result for Males:} \quad P_{50\%} = 187; \text{ or } 273\% \text{ increase}$$

$$\text{Result for Females:} \quad P_{50\%} = 127; \text{ or } 154\% \text{ increase}$$

“BREAK POINT”

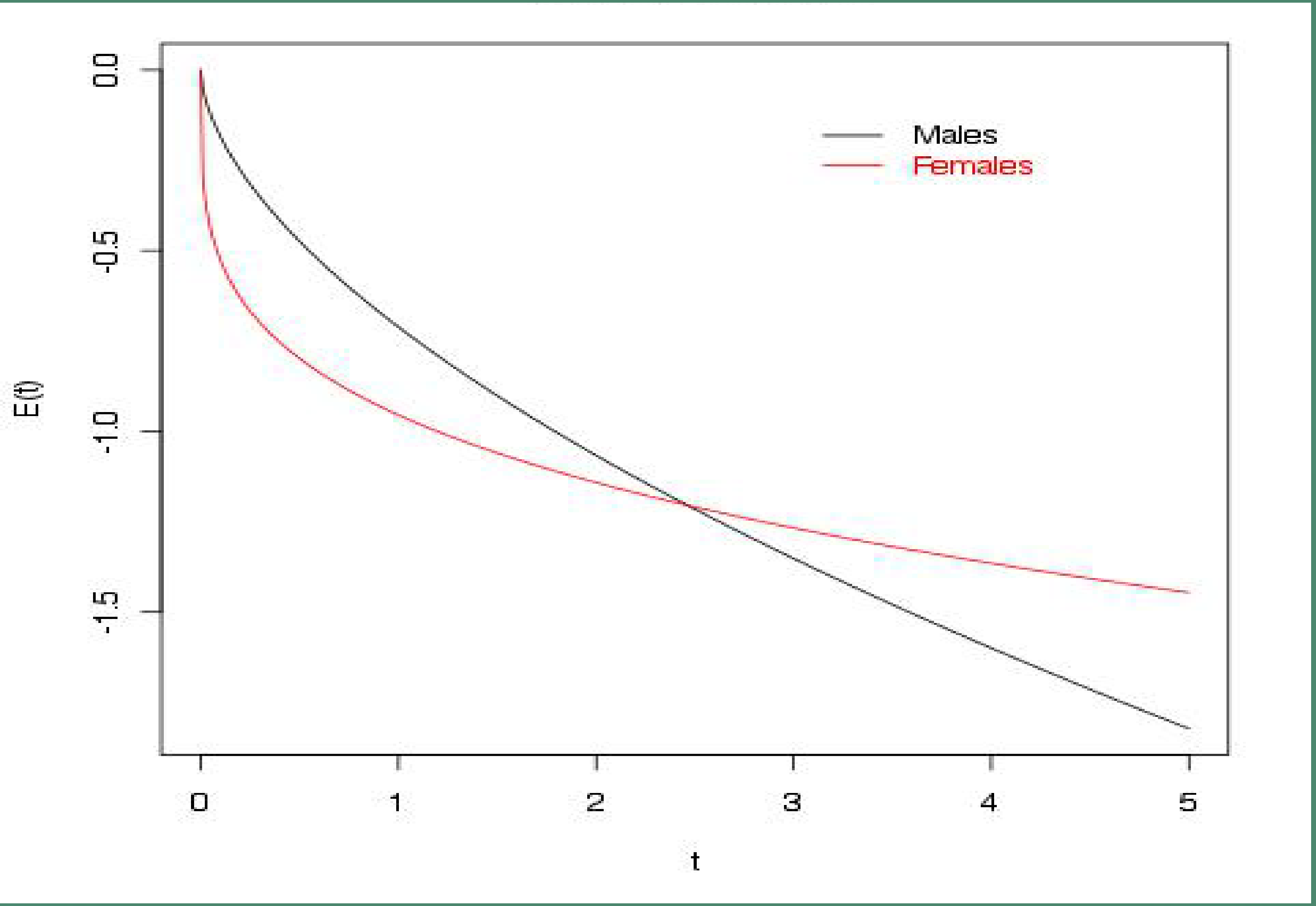
The are under the standardized demand curve is the mean (expected value) of the quantity on the horizontal axis (i.e. log of P/P_0); from this and the Weibull model, we can solve for the Breakpoint (the first price at which consumption is zero), P_{stop} – even individual experiments stop before those breaks. We can obtain numerical solution but, unfortunately, there is no “simple” closed-form solution (it involves the Gamma function).

$$P_{Stop} = P_0 \exp\left\{\frac{\Gamma(1+\frac{1}{\beta})}{\alpha}\right\}$$

$$= 221 \text{ for males}$$

$$= 157 \text{ for females}$$

ELASTICITY CURVES



(1) For both males and females, we have “decreasing elasticity”; consumption reduction accelerates as prices increases – we wonder if this is true for human data;

(2) What’s interesting is the two curves are crossing at a very high price; for lower prices the consumption for females drops faster first but it becomes slower at higher prices.

(3) One possible explanation is that female rats are weaker (larger α, early reduction) but more addictive (smaller β , more resistant to reduction, difference narrows down).

ALTERNATIVE ANALYSIS STRATEGY

1. Adopt either Weibull or Log-logistic model;
2. Going right to population curve, and treat individual data as repeated observations.
3. Use software such as SAS Proc NLMIXED to estimate parameters (& obtain variances)

THE TWO MODELS BY HURSH

There are three levels to characterized how fast the "hazard" is changing: Constant, Weibull (polynomial hazard), and Gompertz (exponential hazard); and we can prove that the Hursh et al. first model (1989) is derivable from the constant hazard (linear elasticity) and the Hursh-Silberberg model (2008) is derivable from the Gompertz Hazard.

THE HURSH-SYLBERBERG MODEL

$$\ln Q = \ln Q_0 + k[e^{-\alpha P} - 1]$$

- P is Price
- Q is Demand/Consumption
- Q_0 is Level of demand when price approaches 0
- k is related to the range of Q
- α is a measure of elasticity

DERIVATION OF THE MODEL

It starts with the Gompertz's Hazard:

$$h(P) = \beta e^{-\alpha P}$$

$$S(P) = \exp[-\int_0^P h(u)du]$$

$$\ln S(P) = -\int_0^P \beta e^{-\alpha u} du$$

$$= \frac{\beta}{\alpha}[e^{-\alpha P} - 1] = \ln \frac{Q}{Q_0}$$

$$\ln Q = \ln Q_0 + \frac{\beta}{\alpha}[e^{-\alpha P} - 1]$$

SCOPE OF THE MODEL

The “target” of investigations using Hursh and Silberberg model (2008) are reinforcers for which the demand curve goes down exponentially (following Gompertz hazard), faster than the constant hazard (Hursh first model, 1989), even faster than the Weibull (which follows a polynomial hazard)

ELASTICITY FUNCTION

$$E(P) = \frac{d[\ln Q]}{d[\ln P]}$$

$$= \frac{d[\ln Q]}{dP}\,\frac{dP}{d[\ln P]}$$

$$= -k\alpha P e^{-\alpha P}$$

$$\lim_{P \to 0} E(P) = 0$$

The system starts at zero elasticity and goes down; Setting E(P) = -1, we get P_{max}

P_{max}

$$k\alpha P_{max}e^{-\alpha P_{max}} = 1$$

At prices below P_{max}, demand or consumption is relatively stable (inelastic or under-elastic); after P_{max} is crossed, consumption becomes over-elastic and falls rapidly with rising price, (that's where addiction starts loosing its grip).

THE CONSTANT k

$$\ln Q = \ln Q_0 + k[e^{-\alpha P} - 1]$$

$$\lim_{P\to\infty} \ln Q = \ln Q_0 - k$$

$$k = \lim_{P\to 0} \ln Q - \lim_{P\to\infty} \ln Q$$

k is the range of lnQ.

Some would argue that k is not estimable because it goes beyond the range of the data. Therefore, many investigators often set it at certain constant level depending on the experiment setting under investigation (Question: Is that "robust" in the estimation of α; if not, how to defend the choice?)

IMPLICATION

$$k\alpha P_{max} e^{-\alpha P_{max}} = 1$$

With k being a constant, α and P_{max} are “inseparable”; their product is constant, one is inversely proportional to the other. In other words, P_{max} is just another measure of elasticity but it’s the one with a much more interesting interpretation than α. Since the product is constant, if α can be normalized, so does P_{max}.

O_{max}

With P the unit price and Q the consumption, the cost or effort (also called output for “O”) is O = PQ

$$\frac{dO}{dP} = Q + P\frac{dQ}{dP}$$

$$= Q + P(Q/P)\,\frac{d[\ln Q]}{d[\ln P]}$$

$$= Q[1 + E(P)]$$

The (total) cost or effort is maximized at P_{max}, when elasticity E(P) = -1

$$O_{max} = (PQ)_{|P=P_{max}}$$
$$= P_{max} Q_0 \exp[k(e^{-\alpha P_{max}} - 1)]$$

O_{max} is another interesting outcome variable to study; unlike P_{max}, it involves elasticity and consumption. It also enriches the interpretation of P_{max}. In animal studies, for example, at Pmax the animal is willing to expend the maximum effort defending Q_0, the freebie; that maximum effort is O_{max}. After P_{max} is crossed, the animal starts to give up; total effort is reduced.

ESTIMATION OF PARAMETERS

We can estimate α (and Q_0) using either Prism or SAS.

Estimates maybe different because of different estimation techniques but, practically, they both are fine and acceptable. And we can obtain measures of precision (variances, standard errors).

After that, P_{max} and O_{max} are calculated from:

$$k\alpha P_{max} e^{-\alpha P_{max}} = 1$$

$$O_{max} = P_{max} Q_0 \exp[k(e^{-\alpha P_{max}} - 1)]$$

There is no closed form solution but we could use Excel; Excel 2010 has a built-in function called SOLVER.

A SIMPLE DATA ANALYSIS

Follow that outlined process, we could calculate individual P_{max} (O_{max}) values, then compare two experimental conditions (or simply males versus females) using either the two-sample or one-sample t-test depending whether the design is unpaired or paired. We could try more fancy method but simple method, such as t-tests, works.