# Urban Travel Demand: A Behavioral Analysis
## Chapters 3, 4, & 5

Tom Domencich and Daniel L. McFadden
North-Holland Publishing Co., 1975

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 312, Spring 2021

# Chapter 3
# A Theory of Individual Travel Demand

$$\operatorname*{Max}_{a \in A} W(Ma, s), \tag{3.1}$$

$$B = \{Na \mid a \in A\},$$

and utility can be written

$$u = U(x, s) \equiv W(MN^{-1}x, s),$$

so that (3.1) becomes

$$\operatorname*{Max}_{x \in B} U(x, s). \tag{3.2}$$

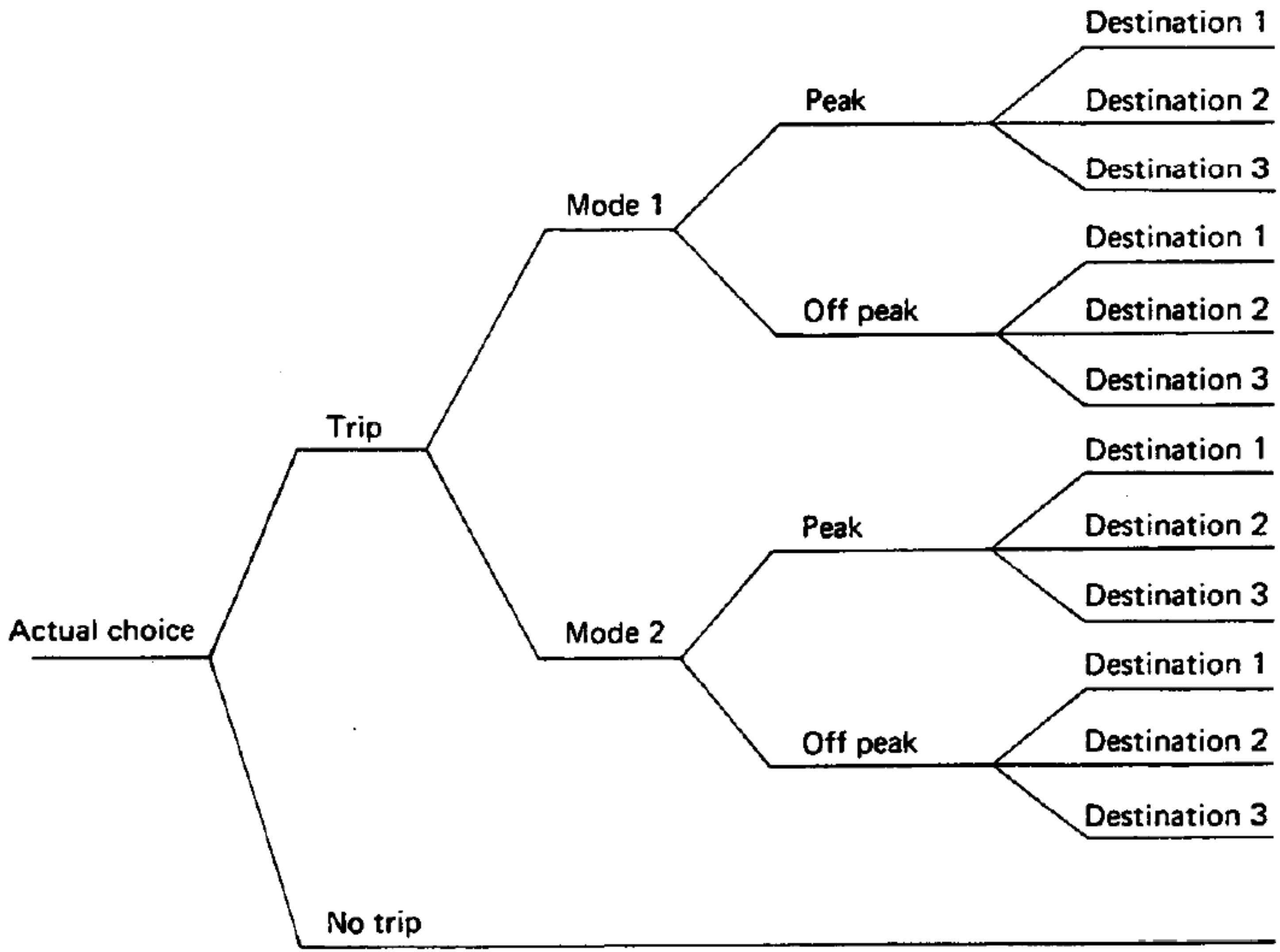

Actual choice
Trip
No trip
Mode 1
Mode 2
Peak
Off peak
Peak
Off peak
Destination 1
Destination 2
Destination 3
Destination 1
Destination 2
Destination 3
Destination 1
Destination 2
Destination 3
Destination 1
Destination 2
Destination 3

Fig. 3.1.

$$U(x^a_{(1)}, x_{(2)}, s) \geqq U(x^b_{(1)}, x_{(2)}, s), \tag{3.3}$$

if and only if

$$U(x^a_{(1)}, x'_{(2)}, s) \geqq U(x^b_{(1)}, x'_{(2)}, s). \tag{3.4}$$

A sufficient condition for this property is that $U$ have the additively separable form[2]

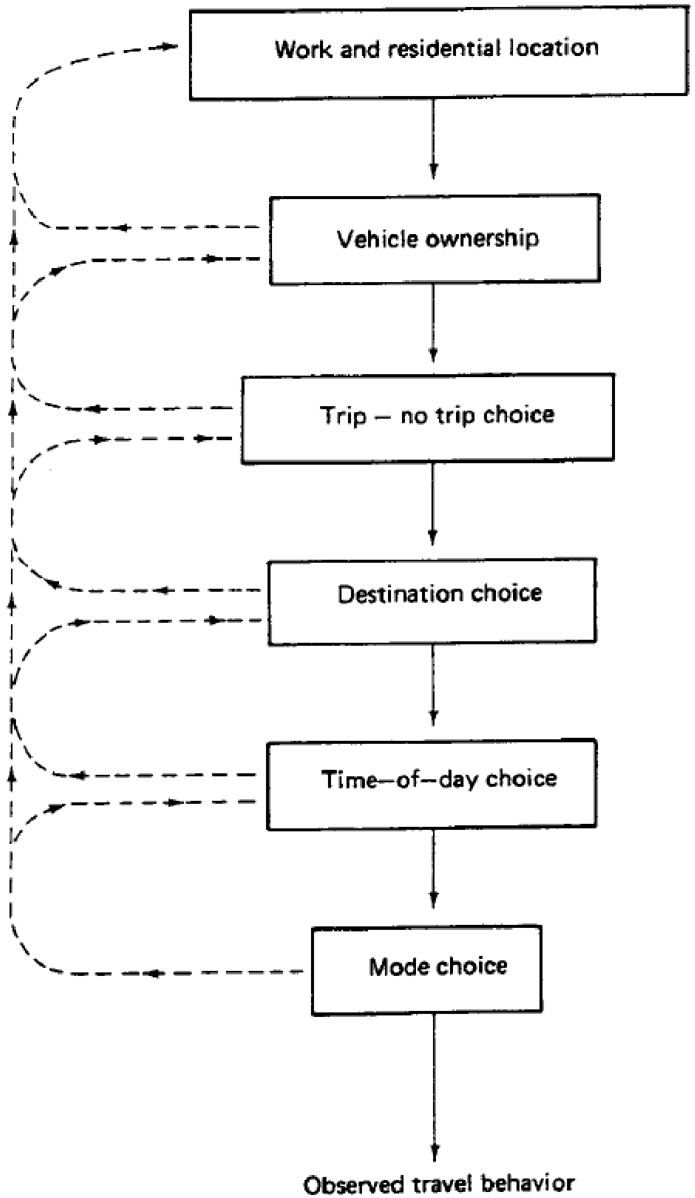
$$U(x_{(1)}, x_{(2)}, s) = \psi(\phi^1(x_{(1)}, s) + \phi^2(x_{(2)}, s)). \tag{3.5}$$

Then, $a$ is chosen over $b$ if

$$\phi^1(x^a_{(1)}, s) > \phi^1(x^b_{(1)}, s). \tag{3.6}$$

$$U(x_{(1)}, x_{(2)}, x_{(3)}, s) = \psi\left(\sum_{i=1}^{3} \phi^i(x_{(i)}, s)\right), \tag{3.7}$$

$$\phi^2(x^p_{(2)}, s) + \underset{i=a,b}{\text{Max}}\, \phi^1(x^{ip}_{(1)}, s) > \phi^2(x^n_{(2)}, s) + \underset{i=a,b}{\text{Max}}\, \phi^1(x^{in}_{(1)}, s). \tag{3.8}$$


Work and residential location
Vehicle ownership
Trip – no trip choice
Destination choice
Time–of–day choice
Mode choice
Observed travel behavior
Fig. 3.2.

$$U(x, s) = \psi\left(\sum_{i=1}^{7} \phi^i(x_{(i)}, s)\right), \tag{3.9}$$

where the subvectors of $x$ are defined as follows:

$x_{(1)}$ Attributes of the trip mode choice, such as fares, tolls, in-vehicle time, walk time, and wait time.

$x_{(2)}$ Attributes of the time-of-day choice, such as congestion at the destination, availability of services, convenience of scheduling other activities.

$x_{(3)}$ Attributes of the destination, such as variety and cost of services available.

$x_{(4)}$ Attributes of the trip—no-trip choice, such as levels of inventories of household goods, characteristics of recreational opportunities at home.

$x_{(5)}$ Attributes of the choice of vehicle ownership, such as the availability, cost, and reliability of household autos.

$x_{(6)}$ Attributes of the choice of residential and work location, such as neighborhood density and quality of public services.

$x_{(7)}$ Attributes of all consumer choices other than the sequence above, including decisions on trips for other purposes and on other days.

# Chapter 4
# A Theory of Population Travel Demand Behavior

$$x = h(B, s; \varepsilon). \tag{4.1}$$

$$u = U(x, s) \tag{4.2}$$

The individual will choose option $i$ if this is the alternative which maximizes his utility; i.e., the individual will choose $i$ if

$$U(x^i, s) > U(x^j, s) \quad \text{for } j \neq i,\ j = 1, \ldots, J. \tag{4.3}$$

Since these utility values are stochastic, the event that the condition in eq. (4.2) holds will occur with some probability, which we can denote by

$$\begin{aligned} P_i &= H(B, s, i) \\ &= \text{Prob}\left[U(x^i, s) > U(x^j, s) \quad \text{for } j \neq i,\ j = 1, \ldots, J\right]. \end{aligned} \tag{4.4}$$

$$U(x, s) = V(x, s) + \eta(x, s), \tag{4.5}$$

where $V$ is non-stochastic and reflects the "representative" tastes of the population, while $\eta$ is stochastic (with mean independent of $x$) and reflects the effect of individual idiosyncrasies in tastes or unobserved attributes for alternatives in $B$. Then eq. (4.3) can be written as

$$P_i = \text{Prob}\left[\eta(x^j, s) - \eta(x^i, s) < V(x^i, s) - V(x^j, s) \text{ for } j \neq i,\ j = 1, \ldots, J\right]. \tag{4.6}$$

Let $\psi(t_1, \ldots, t_J)$ denote the cumulative joint distribution function of $(\eta(x^1, s), \ldots, \eta(x^J, s))$. Let $\psi_i$ denote the derivative of $\psi$ with respect to its $i$th argument, and let $V_j = V(x^j, s)$. Then, eq. (4.6) becomes

$$P_i = \int_{-\infty}^{+\infty} \psi_i(t + V_i - V_1, \ldots, t + V_i - V_J)\mathrm{d}t. \tag{4.7}$$

We consider an individual with a choice between two alternatives (indexed $j = 1, 2$), with vectors of attributes $x^1$ and $x^2$, respectively. The choice probability for the first alternative is then given by eq. (4.7) as

$$P_1 = \int_{-\infty}^{+\infty} \psi_1(t, t + V(x^1, s) - V(x^2, s))\mathrm{d}t, \tag{4.8}$$

$$P_1 = G(V(x^1, s) - V(x^2, s)). \tag{4.9}$$

The form of the functions $V$ and $G$ will be influenced both by the implications of our theory of individual choice behavior and by the constraints of computational practicality.

For the purposes of this discussion (and in our empirical analysis), we assume that $V$ has the general form

$$V(x, s) = Z^1(x, s)\beta_1 + \dots + Z^k(x, s)\beta_k = Z(x, s)'\beta, \quad (4.10)$$

If the distribution function $G$ is linear over the range of $V$, then eqs. (4.9) and (4.10) yield

$$P_1 = (Z(x^1, s) - Z(x^2, s))'\beta, \tag{4.11}$$

$$P_1 = \beta_1 T_a - \beta_2 T_b + \beta_3 I + \beta_4. \tag{4.12}$$

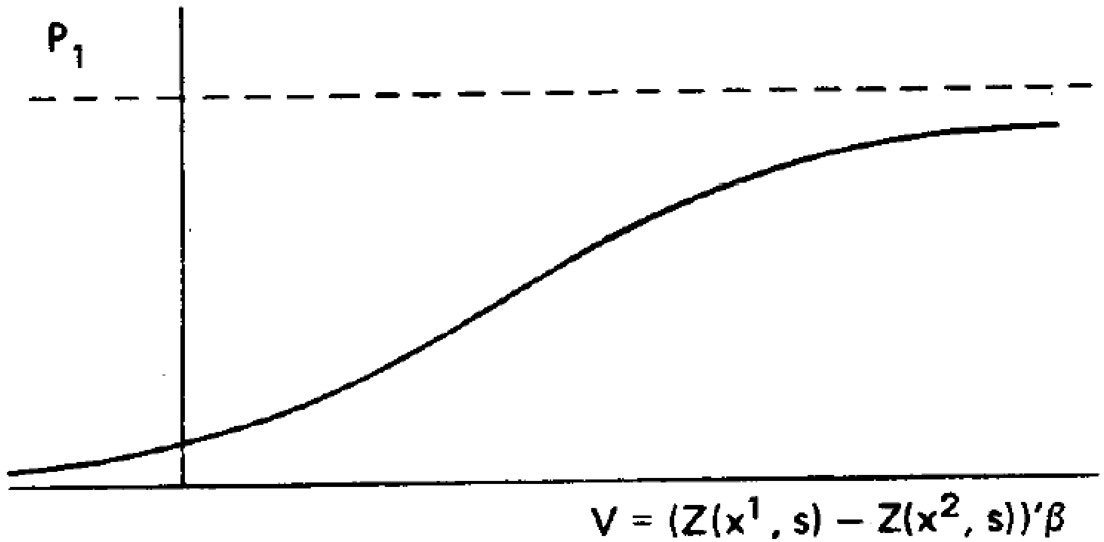



Fig. 4.1. Cumulative probability distributions giving a two-tailed ogive curve.

Normal: $G(V) = \Phi(V) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{V} e^{-t^2/2}\,dt$

Logistic: $G(V) = \frac{1}{1 + e^{-V}}$

Arctan: $G(V) = \frac{1}{\pi} \tan^{-1}(V) + \frac{1}{2}$

$$P_1 = \Phi(\beta' Z(x^1, s) - \beta' Z(x^2, s)), \tag{4.13}$$

where $\Phi$ is the standard cumulative normal distribution. This equation is termed the *binary probit probability model*. The Cauchy distribution gives the probability function

$$P_1 = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\beta' Z(x^2, s) - \beta' Z(x^1, s)), \tag{4.14}$$

and is known as the *arctan probability model*. The logistic distribution gives the probability function

$$P_1 = \frac{1}{1 + \exp\left[\beta' Z(x^2, s) - \beta' Z(x^1, s)\right]}, \qquad (4.15)$$

and is termed the *binary logit probability model.*

$$\log\left[\frac{P_1}{1 - P_1}\right] = \beta_1 T_a - \beta_2 T_b + \beta_3 I + \beta_4. \qquad (4.16)$$

Here, $\beta_1$ and $\beta_2$ measure the change in the log of the *odds* of choosing

Table 4.1

| Argument | Logit[a] | Probit | Arctan[b] | Logit–Probit | Logit–Arctan |
|---|---|---|---|---|---|
| 0.0 | 0.5 | 0.5 | 0.5 | 0.0 | 0.0 |
| 0.2 | 0.5791 | 0.5792 | 0.5781 | −0.0001 | 0.0009 |
| 0.4 | 0.6543 | 0.6554 | 0.6479 | −0.0010 | 0.0064 |
| 0.6 | 0.7226 | 0.7257 | 0.7052 | −0.0031 | 0.0173 |
| 0.8 | 0.7818 | 0.7881 | 0.7504 | −0.0062 | 0.0314 |
| 1.0 | 0.8314 | 0.8413 | 0.7856 | −0.0099 | 0.0457 |
| 1.2 | 0.8715 | 0.8849 | 0.8132 | −0.0133 | 0.0583 |
| 1.4 | 0.9032 | 0.9192 | 0.8351 | −0.0159 | 0.0681 |
| 1.6 | 0.9277 | 0.9451 | 0.8527 | −0.0174 | 0.0750 |
| 1.8 | 0.9464 | 0.9640 | 0.8671 | −0.0176 | 0.0792 |
| 2.0 | 0.9605 | 0.9772 | 0.8791 | −0.0167 | 0.0813 |
| 2.2 | 0.9709 | 0.9860 | 0.8892 | −0.0151 | 0.0817 |
| 2.4 | 0.9787 | 0.9918 | 0.8978 | −0.0130 | 0.0809 |
| 2.6 | 0.9844 | 0.9953 | 0.9052 | −0.0108 | 0.0792 |
| 2.8 | 0.9886 | 0.9974 | 0.9116 | −0.0087 | 0.0770 |
| 3.0 | 0.9917 | 0.9986 | 0.9172 | −0.0069 | 0.0744 |
| 4.0 | 0.9983 | 1.0 | 0.9373 | −0.0016 | 0.0609 |
| 5.0 | 0.9996 | 1.0 | 0.9496 | −0.0003 | 0.0500 |
| 6.0 | 0.9999 | 1.0 | 0.9579 | 0.0 | 0.0420 |
| 7.0 | 1.0 | 1.0 | 0.9638 | 0.0 | 0.0361 |
| 8.0 | 1.0 | 1.0 | 0.9683 | 0.0 | 0.0316 |
| 9.0 | 1.0 | i.0 | 0.9718 | 0.0 | 0.0281 |
| 10.0 | 1.0 | 1.0 | 0.9746 | 0.0 | 0.0253 |

[a] The logit formula (normalized to have the same slope at zero as the standard normal) is $P = 1/\{1 + \exp[-2x\sqrt{(2/\pi)}]\}$.

[b] The arctan formula (normalized to have the same slope at zero as the standard normal) is $P = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}[x\sqrt{(\pi/2)}]$.

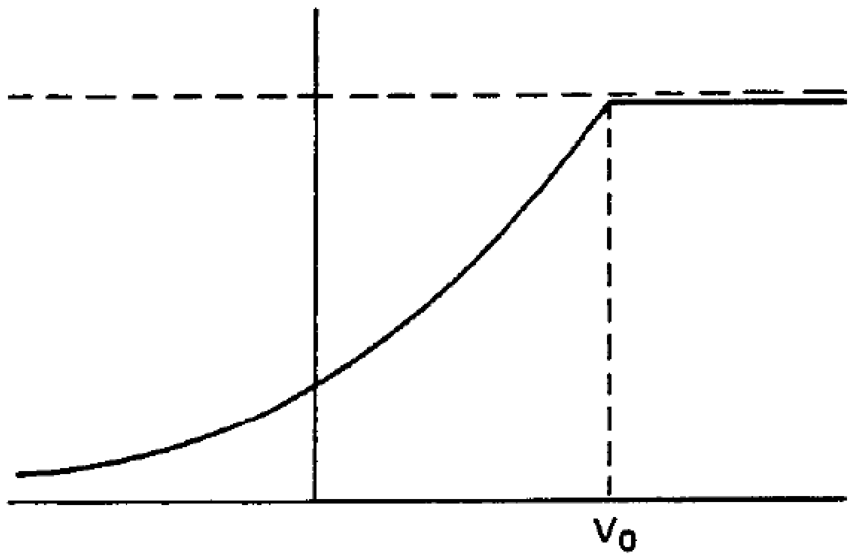



Fig. 4.2. Cumulative exponential probability distribution giving a one-tailed ogive curve. $G(V) = e^{-(V_0 - V)}$ for $V < V_0$; $G(V) = 1$ for $V \geqslant V_0$.

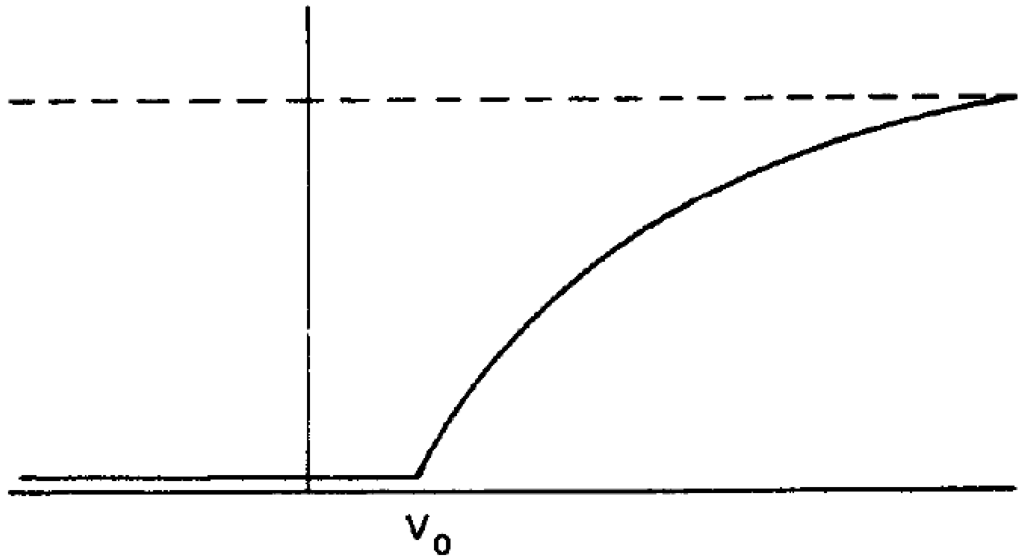



Fig. 4.3. Cumulative exponential probability distribution giving a one-tailed ogive curve. $G(V) = 1 - e^{-(V-V_0)}$ for $V > V_0$; $G(V) = 0$ for $V \leqslant V_0$.

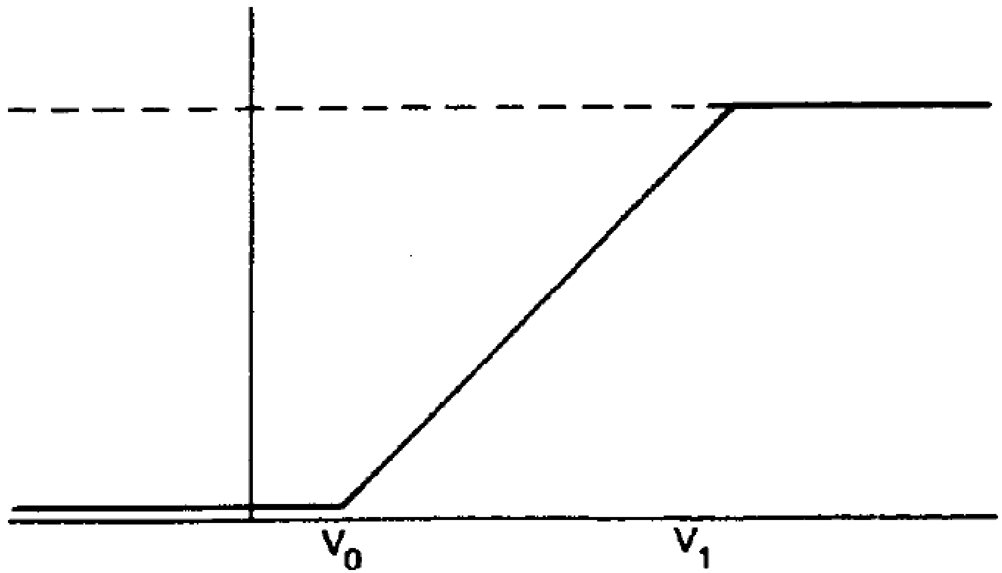



Fig. 4.4. Uniform distribution giving a "truncated linear model". $G(V) = 1$ for $V \geqslant V_1$; $G(V) = (V - V_0)/(V_1 - V_0)$ for $V_0 < V < V_1$; $G(V) = 0$ for $V < V_0$.

$$G(v) = \int_{-\infty}^{+\infty} \psi'(t)\psi(v + t)\,\mathrm{d}t. \tag{4.17}$$

A random variable $\eta_i$ has a Weibull (extreme value, Gnedenko) distribution if

$$\mathrm{Prob}\,[\eta_i \leqq \eta] = \mathrm{e}^{-\mathrm{e}^{-(\eta+\alpha)}}, \tag{4.18}$$

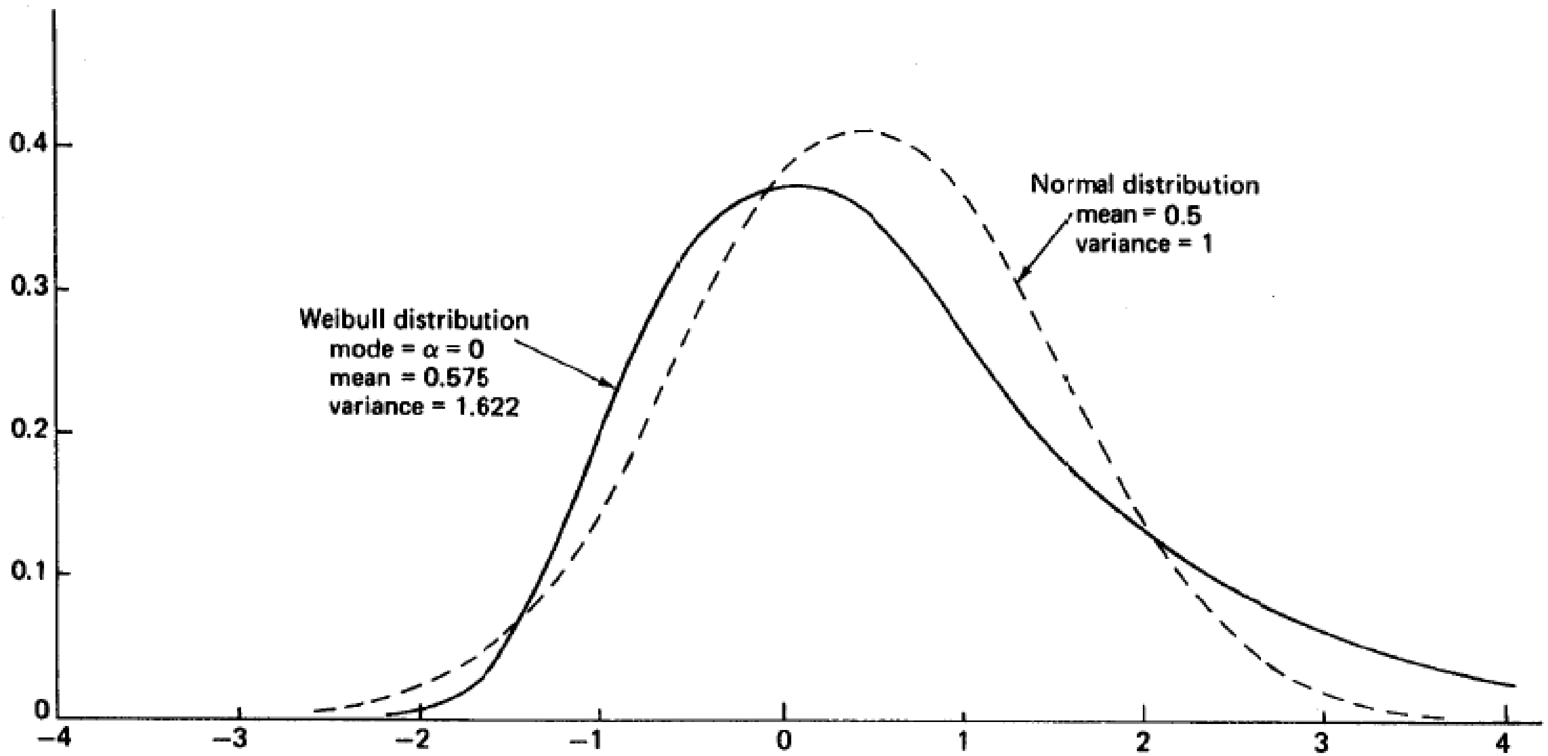



Fig. 4.5. Frequency functions of normal and Weibull distributions.

*Lemma.* If random variables $\eta_i$ have independent Weibull distributions with parameters $\alpha_i$ for $i = 1, \ldots, n$, then:

(a) $\eta_i + v$ has a Weibull distribution with parameter $\alpha_i - v$ for any real $v$;

(b) $\operatorname{Max}_{i=1,\ldots,n} \eta_i$ has a Weibull distribution with parameter

$-\log \sum_{i=1}^{n} e^{-\alpha_i}$;

(c) $$\text{Prob}\,[v_1 + \eta_1 \geqq v_2 + \eta_2] = \frac{e^{v_1 - \alpha_1}}{e^{v_1 - \alpha_1} + e^{v_2 - \alpha_2}}; \tag{4.19}$$

(d) $$\text{Prob}\,[v_1 + \eta_1 \geqq v_i + \eta_i \quad \text{for } i = 2, \ldots, n] = \frac{e^{v_1 - \alpha_1}}{\sum_{i=1}^{n} e^{v_i - \alpha_i}} \tag{4.20}$$

Verification of this result is straightforward; we outline the steps for the sake of completeness. To show (a), note that

$$\text{Prob}\,[\eta_i + v \leqq \eta] = \text{Prob}\,[\eta_i \leqq \eta - v],$$

and substitute the argument $\eta - v$ in the cumulative distribution function of $\eta_i$. To show (b), note that

$$\text{Prob}\,[\max_{i=1,\ldots,n} \eta_i \leqq \eta] = \text{Prob}\,[\eta_1 \leqq \eta] \cdot \ldots \cdot \text{Prob}\,[\eta_n \leqq \eta]$$

$$= \exp\left[-\sum_{i=1}^{n} e^{-(\eta+\alpha_i)}\right]$$

$$= \exp\left[-e^{-\eta} \cdot \sum_{i=1}^{n} e^{-\alpha_i}\right]. \qquad (4.21)$$

Setting

$$e^{-\alpha} = \sum_{i=1}^{n} e^{-\alpha_i}$$

establishes that the maximum value is distributed Weibull with parameter $\alpha$.

To establish (c), we use the convolution formula

$$\text{Prob}\,[v_1 + \eta_1 \geqq v_2 + \eta_2] = \int_{-\infty}^{+\infty} \psi_1'(\eta)\psi_2(v_1 - v_2 + \eta)\,d\eta, \tag{4.22}$$

where $\psi_i$ is the cumulative distribution function of $\eta_i$. In this case

$$\psi_1(\eta) = \exp\,(-e^{-(\eta+\alpha_1)}),$$

$$\psi_1'(\eta) = e^{-(n+\alpha_1)}\exp(-e^{-(n+\alpha_1)}).$$

Then eq. (4.22) becomes

$$\begin{aligned}
&\text{Prob}\,[v_1 + \eta_1 \geqq v_2 + \eta_2] \\
&\quad = \int_{-\infty}^{+\infty} e^{-(\eta+\alpha_1)}\exp(-e^{-(\eta+\alpha_1)})\exp(-e^{-(\eta+v_1-v_2+\alpha_2)})\,d\eta \\
&\quad = \int_{-\infty}^{+\infty} e^{-(\eta+\alpha_1)}\exp(-e^{-\eta}(e^{-\alpha_1} + e^{-v_1+v_2-\alpha_2}))\,d\eta \\
&\quad = \frac{e^{-\alpha_1}}{[e^{-\alpha_1} + e^{-v_1+v_2-\alpha_2}]} \\
&\qquad \cdot \int_0^1 d\,\{\exp(-e^{-\eta}(e^{-\alpha_1} + e^{-v_1+v_2-\alpha_2}))\} \\
&\quad = \frac{e^{-v_1-\alpha_1}}{e^{v_1-\alpha_1} + e^{v_2-\alpha_2}}.
\end{aligned} \tag{4.23}$$

$$\text{Prob}\,[\eta_1 + v_1 \geqq \eta_i + v_i \quad \text{for } i = 1, \ldots, n] =$$

$$\text{Prob}\,[\eta_1 + v_1 \geqq \underset{i=2,\ldots,n}{\text{Max}} (\eta_i + v_i)] = \frac{e^{v_1 - \alpha_1}}{e^{v_1 - \alpha_1} + e^{-\alpha}}, \qquad (4.24)$$

by (c), where

$$\alpha = -\log \sum_{i=2}^{n} e^{v_i - \alpha_i}$$

is the parameter of the Weibull distributed variable $\text{Max}_{i=2,\ldots,n}(\eta_i + v_i)$

As indicated earlier, the importance of this lemma lies in the fact, established in condition (c), that the difference of two independent Weibull distributed random variables has a binary logit distribution

$$\text{Prob}\,[\eta_2 - \eta_1 \leqq v_1 - v_2] \equiv G(v_1 - v_2) = \frac{e^{v_1 - \alpha_1}}{e^{v_1 - \alpha_1} + e^{v_2 - \alpha_2}}. \tag{4.25}$$

$$
\begin{aligned}
P_i = H(B, s, i) &= \text{Prob}\,[U(x^i, s) > U(x^j, s) \\
&\qquad \text{for } j \neq i, j = 1, \ldots, J] \\
&= \int_{-\infty}^{+\infty} \psi_i(t + V_i - V_1, \ldots, t + V_i - V_J)\,\mathrm{d}t, \quad (4.26)
\end{aligned}
$$

$$P_1 = \int_{r_2=-\infty}^{V_1-V_2} \cdots \int_{r_J=-\infty}^{V_1-V_J} n(r;0;\Omega)\mathrm{d}r_2 \ldots \mathrm{d}r_J, \tag{4.27}$$

where $n(r;0;\Omega)$ is the multivariate normal frequency function with mean vector 0 and covariance matrix $\Omega$ evaluated at argument $r$, and where the elements $\omega_{ij}$ of $\Omega$ satisfy

$$\omega_{ij} = E\eta_i\eta_j + E\eta_1^2 - E\eta_i\eta_1 - E\eta_j\eta_1, \tag{4.28}$$

with $\eta_i = U(x_i, s) - V(x_i, s)$ and $E\eta_i = 0$. A further simplification of this formula can be obtained by defining

$$\lambda_{2j} = \omega_{2j}, \qquad t_2 = r_2 \tag{4.29}$$

and recursively, for $i = 3, \ldots, J$,

$$\lambda_{ik} = \omega_{ik} - \sum_{j=2}^{i-1} \lambda_{ji}\lambda_{jk}/\lambda_{jj}, \quad k = 2, \ldots, J, \tag{4.30}$$

$$t_i = r_i - \sum_{j=2}^{i-1} \lambda_{ji}r_j/\lambda_{jj}. \tag{4.31}$$

$$P_1 = \int_{t_2=-\infty}^{V_1 - V_2} \phi\left[\frac{t_2}{\sqrt{\lambda_{22}}}\right] \int_{t_3=-\infty}^{V_1 - V_3 - \frac{\lambda_{23}}{\lambda_{22}} t_2} \phi\left[\frac{t_3}{\sqrt{\lambda_{33}}}\right]$$

$$\int_{t_4=-\infty}^{V_1 - V_4 - \frac{\lambda_{24}}{\lambda_{22}} t_2 - \frac{\lambda_{34}}{\lambda_{33}} t_3} \phi\left[\frac{t_4}{\sqrt{\lambda_{44}}}\right] \cdots \int_{t_J=-\infty}^{V_1 - V_J - \sum_{j=2}^{J-1} \frac{\lambda_{jJ}}{\lambda_{jj}} t_j} \phi\left[\frac{t_J}{\sqrt{\lambda_{JJ}}}\right] \mathrm{d}t_J \ldots \mathrm{d}t_2, \tag{4.32}$$

$$U(x, s) = \sum_{k=1}^{K} \alpha_k Z^k(x, s) + \varepsilon(x, s), \tag{4.33}$$

$$\eta_j = \eta(x^j, s) = \sum_{k=1}^{K} (\alpha_k - \beta_k) Z^k(x^j, s) + \varepsilon_j, \tag{4.34}$$

$$\omega_{jl} = \operatorname{cov}(\eta_j, \eta_l) = \operatorname{cov}(\varepsilon_j, \varepsilon_l) + \sum_{k=1}^{K} \operatorname{var}(\alpha_k) Z^k(x^j, s)^2. \tag{4.35}$$

$$P_1 = \Phi\left[\frac{\beta' Z(x^1, s) - \beta' Z(x^2, s)}{\sqrt{(\omega_{11} + \omega_{22} - 2\omega_{12})}}\right]. \tag{4.36}$$

We next consider the case in which the stochastic components of utility have independent Cauchy distributions. Then,

$$\psi(t_1, \ldots, t_J) = \prod_{j=1}^{J} \left( \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(t_j/a_j) \right), \tag{4.37}$$

where the $a_j$ are positive constants, and eq. (4.26) becomes, for $P_1$,

$$P_1 = \int_{t=-\infty}^{+\infty} \frac{a_1}{a_1^2 + t^2} \prod_{j=2}^{J} \left[ \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left( \frac{V_1 - V_j + t}{a_j} \right) \right] \mathrm{d}t. \tag{4.38}$$

$$P_i = \frac{e^{V_i - \alpha_i}}{\sum_{j=1}^{J} e^{V_j - \alpha_j}}, \quad (4.39)$$

$$P_i = \frac{e^{V(x^i,s)}}{\sum_{j=1}^{J} e^{V(x^j,s)}}. \quad (4.39a)$$

$$\mathrm{Log}(P_i/P_j) = V(x^i, s) - V(x^j, s). \quad (4.40)$$

Given the fitted function $V(x^j, s)$ for $j = 1, 2$, we have the probability function

$$P_i = \frac{e^{V(x^i,s)}}{e^{V(x^1,s)} + e^{V(x^2,s)}} \quad \text{for } i = 1, 2. \tag{4.41}$$

$$P_i = \frac{e^{V(x^i,s)}}{e^{V(x^1,s)} + e^{V(x^2,s)} + e^{V(x^3,s)}}, \tag{4.42}$$

$$
\begin{bmatrix}\text{Percentage} \\ \text{change in } P_1\end{bmatrix} = \frac{P_1(3\text{ modes}) - P_1(2\text{ modes})}{P_1(2\text{ modes})}
$$

$$
= \frac{e^{V_1}/(e^{V_1} + e^{V_2} + e^{V_3})}{e^{V_1}/(e^{V_1} + e^{V_2})} - 1
$$

$$
= \frac{e^{V_1} + e^{V_2}}{e^{V_1} + e^{V_2} + e^{V_3}} - 1
$$

$$
= \frac{e^{V_3}}{e^{V_1} + e^{V_2} + e^{V_3}}
$$

$$
= -P_3, \quad \text{where } V_i = V(x^i, s). \tag{4.43}
$$

$$\begin{bmatrix}\text{Probability of} \\ \text{choosing time } t, \\ \text{mode } m\end{bmatrix} = \begin{bmatrix}\text{Probability of choosing} \\ \text{mode } m\text{, conditioned} \\ \text{on the event that time } t \\ \text{is chosen}\end{bmatrix} \cdot \begin{bmatrix}\text{Probability of} \\ \text{choosing time } t, \\ \text{any mode}\end{bmatrix}. \tag{4.44}$$

$$\begin{bmatrix}\text{Probability of choosing} \\ \text{mode } m\text{, conditioned} \\ \text{on the event that time } t \\ \text{is chosen}\end{bmatrix} = \begin{bmatrix}\text{Probability of choosing} \\ \text{mode } m \text{ when the set of} \\ \text{alternatives is the set of} \\ \text{modes available at time } t\end{bmatrix}. \tag{4.45}$$

$$\begin{bmatrix}\text{Probability} \\ \text{of choosing} \\ \text{time } t\text{, any mode}\end{bmatrix} = \begin{matrix}\text{Sum over modes} \\ m' \text{ available at} \\ \text{time } t\end{matrix} \begin{bmatrix}\text{Probability of} \\ \text{choosing available} \\ \text{mode } m' \text{ at time } t\end{bmatrix}. \tag{4.46}$$

$$P_i = e^{V_i} \Big/ \sum_{j=1}^{J} e^{V_j}$$

$$= \left[ e^{V_i} \Big/ \sum_{j=1}^{i} e^{V_j} \right] \left[ \sum_{j=1}^{i} e^{V_j} \Big/ \sum_{l=1}^{J} e^{V_l} \right], \tag{4.47}$$

$$U(x, s) = \sum_{i=1}^{7} \phi^i(x_{(i)}, s), \tag{4.48}$$

$$U(x, s) = \sum_{i=1}^{7} \phi^i(x_{(i)}, s) + \eta(x, s). \tag{4.49}$$

$$
\begin{aligned}
P_{m|t} &= \text{Prob}\left[U(x^m, s) > U(x^j, s) \quad \text{for all } j \neq m\right] \\
&= \text{Prob}\left[\phi^1(x^m_{(1)}, s) + \eta(x^m, s)\right. \\
&\qquad \left. > \phi^1(x^j_{(1)}, s) + \eta(x^j, s) \quad \text{for } j \neq m\right],
\end{aligned}
\tag{4.50}
$$

$$
P_{m|t} = \frac{e^{\phi^1(x^m_{(1)}, s) - \alpha_{mt}}}{\sum_j e^{\phi^1(x^m_{(1)}, s) - \alpha_{jt}}},
\tag{4.51}
$$

$$
\begin{aligned}
P_t &= \text{Prob}\,[\operatorname*{Max}_j U(x^{jt}, s) > \operatorname*{Max}_j U(x^{jp}, s) \quad \text{for } p \neq t] \\
&= \text{Prob}\,[\phi^2(x^t_{(2)}, s) + \operatorname*{Max}_j \{\phi^1(x^{jt}_{(1)}, s) + \eta(x^{jt}, s)\} > \\
&\qquad \phi^2(x^p_{(2)}, s) + \operatorname*{Max}_j \{\phi^1(x^{jp}_{(1)}, s) + \eta(x^{jp}, s)\} \quad \text{for } p \neq t].
\end{aligned} \tag{4.52}
$$

Provided that the $\eta(x^{jp}, s)$ have independent Weibull distributions for each index $(j, p)$, we conclude that $\eta_p$, defined by

$$\operatorname*{Max}_j \phi^1(x^{jp}_{(1)}, s) + \eta_p = \operatorname*{Max}_j \{\phi^1(x^{jp}_{(1)}, s) + \eta(x^{jp}, s)\}, \qquad (4.53)$$

$$P_t = \text{Prob}\,[\phi^2(x^t_{(2)}, s) + \underset{j}{\text{Max}}\,\phi^1(x^{jt}_{(1)}, s) + \eta_t >$$

$$\phi^2(x^p_{(2)}, s) + \underset{j}{\text{Max}}\,\phi^1(x^{jp}_{(1)}, s) + \eta_p \quad \text{for } p \neq t] \tag{4.54}$$

$$= \frac{\exp\,[\phi^2(x^t_{(2)}, s) + \underset{j}{\text{Max}}\,\phi^1(x^{jt}_{(1)}, s) - a_t]}{\sum_p \exp\,[\phi^2(x^p_{(2)}, s) + \underset{j}{\text{Max}}\,\phi^1(x^{jp}_{(1)}, s) - a_p]}, \tag{4.55}$$

where the Weibull distribution "parameters" $a_p$ satisfy (using Lemma conclusions (a) and (b))

$$a_p = \underset{p}{\text{Max}}\,\phi^1(x^{jp}_{(1)}, s) - \log \sum_j \exp\,[-\alpha_{jp} + \phi^1(x^{jp}_{(1)}, s)]. \tag{4.56}$$

Substituting this expression in eq. (4.55) then yields

$$P_t = \frac{\exp[\phi^2(x^t_{(2)}, s)] \sum_j \exp[-\alpha_{jt} + \phi^1(x^{jt}_{(1)}, s)]}{\sum_p \exp[\phi^2(x^p_{(2)}, s)] \sum_j \exp[-\alpha_{jt} + \phi^1(x^{jt}_{(1)}, s)]}$$

$$= \frac{\sum_j \exp[\phi^1(x^{jt}_{(1)}, s) + \phi^2(x^t_{(2)}, s) - \alpha_{jt}]}{\sum_p \sum_j \exp[\phi^1(x^{jp}_{(1)}, s) + \phi^2(x^p_{(2)}, s) - \alpha_{jp}]}. \tag{4.57}$$

Combining eqs. (4.51) and (4.57) in the formula (4.44) yields the choice probability $P_{mt}$ for the simultaneous choice of mode $m$ and time $t$,

$$
\begin{aligned}
P_{mt} &= P_{m|t}P_t \\
&= \frac{e^{\phi^1(x^{mt}_{(1)},s) - \alpha_{mt}}}{\sum_j e^{\phi^1(x^{mt}_{(1)},s) - \alpha_{jt}}} \frac{\sum_j e^{\phi^1(x^{jt}_{(1)},s) + \phi^2(x^t_{(2)},s) - \alpha_{jt}}}{\sum_p \sum_j e^{\phi^1(x^{jp}_{(1)},s) + \phi^2(x^p_{(2)},s) - \alpha_{jp}}} \\
&= \frac{e^{\phi^1(x^{mt}_{(1)},s) + \phi^2(x^t_{(2)},s) - \alpha_{mt}}}{\sum_p \sum_j e^{\phi^1(x^{jp}_{(1)},s) + \phi^2(x^p_{(2)},s) - \alpha_{jp}}} .
\end{aligned}
\tag{4.58}
$$

$$
y_p = -\log \sum_j \exp\left[-\alpha_{jp} + \phi^1(x^{jp}_{(1)}, s)\right] \tag{4.59}
$$

$$\phi^1(x^{jp}_{(1)}, s) = \sum_{k=1}^{K} \beta_k Z^k(x^{jp}_{(1)}, s). \tag{4.60}$$

Define

$$q_j = -\alpha_{jp} + \sum_{k=1}^{K} \beta_k Z^k(x^{jp}_{(1)}, s),$$

and

$$y_p = y(q_1, \ldots, q_J) = -\log \sum_j e^{q_j}$$

Let $\bar{q}$ denote the average of the $q_j$, and make a first-order Taylor's expansion of $y(\bar{q}, \ldots, \bar{q})$ about the vector of values $(q_1, \ldots, q_J)$:

$$\begin{aligned} y(\bar{q}, \ldots, \bar{q}) &\equiv -\log J - \bar{q} \\ &= y(q_1, \ldots, q_J) + \sum_j \frac{\partial y}{\partial q_j}\bigg|_{(q_1, \ldots, q_J)} \cdot (\bar{q} - q_j) \\ &\quad + \text{higher-order terms in } (\bar{q} - q_j), \end{aligned} \tag{4.61}$$

but

$$\frac{\partial y}{\partial q_i} = \frac{-e^{q_i}}{\sum_j e^{q_j}}.$$

Comparing this formula with eq. (4.51),

$$\frac{\partial y}{\partial q_i} = -P_{i|p}.$$

Hence, from eq. (4.61),

$$\begin{aligned} y_p = y(q_1, \ldots, q_J) &= y(\bar{q}, \ldots, \bar{q}) - \sum_j \frac{\partial y}{\partial q_j}\bigg|_{(q_1, \ldots, q_J)} \cdot (\bar{q} - q_j) \\ &\quad - \text{higher-order terms} \\ &= -\log J - \bar{q} + \sum_j P_{j|p} \cdot \bar{q} - \sum_j P_{j|p} q_j \\ &\quad - \text{higher-order terms} \\ &= -\log J + \sum_j P_{j|p} \alpha_{jp} - \sum_{k=1}^{K} \beta_k \sum_j P_{j|p} Z^k(x_{(1)}^{jp}, s) \\ &\quad - \text{higher-order terms in } (q_j - \bar{q}). \end{aligned}$$

$$y_p = -\sum_{k=1}^{K} \beta_k \left[ \sum_{j=1}^{J} P_{j|p} Z^k(x_{(1)}^{jp}, s) \right], \tag{4.62}$$

and

$$P_t = \frac{\exp\left[\phi^2(x_{(2)}^t, s) - y_t\right]}{\sum_p \exp\left[\phi^2(x_{(2)}^p, s) - y_p\right]}. \tag{4.63}$$

$$V(x^i, s) = \log\left[\sum_{k=1}^{K} \beta_k Z^k(x^i, s)\right]. \tag{4.64}$$

Suppose further that

$$0 \leqq \sum_{k=1}^{K} \beta_k Z^k(x^i, s) \leqq 1 \quad \text{for } i = 1, \ldots, J,$$

and

$$\sum_{i=1}^{J} \sum_{k=1}^{K} \beta_k Z^k(x^i, s) = 1. \tag{4.65}$$

Then the probability functions have the form

$$P_i = h(B, s, i) = \sum_{k=1}^{K} \beta_k Z^k(x^i, s). \tag{4.66}$$

The second concrete specification we consider for the $V(x, s)$ function directly incorporates the restrictions above, and leads to a multinomial extension of the logit model. Suppose $V$ has the linear-in-parameters form

$$V(x^i, s) = \sum_{k=1}^{K} \beta_k Z^k(x^i, s). \tag{4.67}$$

Then, from eq. (4.39), the probability functions have the form

$$P_i = h(B, s, i) = \frac{1}{\sum_{l=1}^{J} \exp \sum_{k=1}^{K} \beta_k [Z^k(x^l, s) - Z^k(x^i, s)]}. \tag{4.68}$$

$$D_j = \sum_{i=1}^{I} N_i P_j^i. \tag{4.69}$$

$$D_j = \sum_{i=1}^{I} P_j^i \theta_i, \tag{4.70}$$

$$D_1 = \frac{N}{\sqrt{(\beta'\Omega\beta)}} \int_{-\infty}^{+\infty} \Phi(t)\, \phi\left[\frac{t - \beta'\bar{z}}{\sqrt{(\beta'\Omega\beta)}}\right] \mathrm{d}t. \tag{4.71}$$

If $X_k$ is normal with mean $\mu_k$ and variance $\sigma_k^2$ for $k = 1, 2$, then, utilizing the convolution property of two normal random variables,

$$\begin{aligned} \operatorname{Prob}[X_1 - X_2 \leqq x] &= \Phi\left[\frac{x - \mu_1 + \mu_2}{\sqrt{(\sigma_1^2 + \sigma_2^2)}}\right] \\ &= \frac{1}{\sigma_2} \int_{-\infty}^{+\infty} \Phi\left[\frac{t + x - \mu_1}{\sigma_1}\right] \phi\left[\frac{t - \mu_2}{\sigma_2}\right] \mathrm{d}t. \end{aligned} \tag{4.72}$$

$$D_1 = N\Phi\left[\frac{\beta'\bar{z}}{\sqrt{(1+\beta'\Omega\beta)}}\right]. \tag{4.73}$$

$$P_j^i = \frac{e^{\beta' z^{ji}}}{\sum_{l=1}^{J_i} e^{\beta' z^{li}}}, \tag{4.73a}$$

$$\frac{\partial(N_i P_j^i)}{\partial z_k^{ji}} = \beta_k N_i P_j^i (1 - P_j^i). \tag{4.73b}$$

$$\frac{\partial(N_i P_j^i)}{\partial z_k^{li}} = -\beta_k N_i P_j^i P_l^i, \quad l \neq j. \tag{4.73c}$$

$$E_j^i(j,k) \equiv \frac{z_k^{ji}}{N_i P_j^i} \frac{\partial (N_i P_j^i)}{\partial z_k^{ji}} = \beta_k z_k^{ji}(1 - P_j^i). \tag{4.73d}$$

$$E_j^i(l,k) = \frac{z_k^{li}}{N_i P_j^i} \frac{\partial (N_i P_j^i)}{\partial z_k^{li}} = -\beta_k z_k^{li} P_l^i. \tag{4.73e}$$

$$D_j = \sum_j N_i P_j^i, \tag{4.73f}$$

$$E_j(j,k) \equiv \frac{t}{D_j}\frac{\partial D_j}{\partial t}\bigg|_{t=1} = \left[\frac{t}{\sum_i N_i P_j^i}\sum_i \frac{\partial(N_i P_j^i)}{\partial(\bar{z}_k^{ji}t)}\frac{\partial(\bar{z}_k^{ji}t)}{\partial t}\right]\bigg|_{t=1}$$

$$= \sum_i w_i \frac{\partial(N_i P_j^i)}{\partial(\bar{z}_k^{ji}t)}\bigg|_{t=1}\frac{\bar{z}_k^{ji}}{N_i P_j^i} \tag{4.73g}$$

$$= \sum_i w_i E_j^i(j,k),$$

where

$$w_i = N_i P_j^i \Big/ \sum_l N_l P_j^l$$

$$N = \sum_x \sum_s N(x, s) \tag{4.74}$$

$$T_{mtodps} = \sum_x N(x, s) \sum_{j \in J_{mtodp}} P_j(x, s). \tag{4.75}$$

It is convenient to also distinguish the total number of trips of type *mtodp* taken by each homogeneous subpopulation,

$$T_{mtodp}(x, s) = N(x, s) \sum_{j \in J_{mtodp}} P_j(x, s), \tag{4.76}$$

so that

$$T_{mtodps} = \sum_x T_{mtodp}(x, s). \tag{4.77}$$

$$T_{todp}(x, s) = N(x, s) \sum_{j \in J_{todp}} P_j(x, s)$$

$$= N(x, s) \sum_{m} \sum_{j \in J_{mtodp}} P_j(x, s), \tag{4.78}$$

and

$$T_{todps} = \sum_{x} T_{todp}(x, s). \tag{4.79}$$

$$T_{top}(x, s) = N(x, s) \sum_{j \in J_{top}} P_j(x, s), \tag{4.80}$$

$$T_{tops} = \sum_{x} T_{top}(x, s). \tag{4.81}$$

$$T_{to}(x, s) = N(x, s) \sum_{j \in J_{to}} P_j(x, s) \tag{4.82}$$

As a notational shorthand, write

$$Q_{mtodp} \equiv Q_{mtodp}(x, s) \equiv \sum_{j \in J_{mtodp}} P_j(x, s), \tag{4.83}$$

with analogous definitions for other subscripts. Then

$$\begin{aligned} T_{mtodp}(x, s) &= N(x, s)Q_{mtodp}(x, s) \\ &= \left[\frac{Q_{mtodp}(x, s)}{Q_{todp}(x, s)}\right] \cdot \left[\frac{Q_{todp}(x, s)}{Q_{top}(x, s)}\right] \cdot \left[\frac{Q_{top}(x, s)}{Q_{to}(x, s)}\right] \cdot N(x, s)Q_{to}(x, s) \\ &\equiv Q_{m|todp}(x, s) \cdot Q_{d|top}(x, s) \cdot Q_{p|to}(x, s) \cdot T_{to}(x, s), \end{aligned} \tag{4.84}$$

where we have defined

$$Q_{m|todp}(x, s) = Q_{mtodp}(x, s)/Q_{todp}(x, s), \tag{4.85}$$

From eq. (4.84), we obtain the marginal trip tables for the subpopulation,

$$T_{todp}(x, s) = Q_{d|top}(x, s) Q_{p|to}(x, s) T_{to}(x, s), \tag{4.86}$$

$$T_{top}(x, s) = Q_{p|to}(x, s) T_{to}(x, s). \tag{4.87}$$

Aggregating the formulae (4.84), (4.86) and (4.87) over $x$ provides the basic trip tables of interest:

$$T_{mtodps} = \sum_x Q_{m|todp}(x, s) Q_{d|top}(x, s) Q_{p|to}(x, s) T_{to}(x, s), \tag{4.88}$$

$$T_{todps} = \sum_x Q_{d|top}(x, s) Q_{p|to}(x, s) T_{to}(x, s), \tag{4.89}$$

$$T_{tops} = \sum_x Q_{p|to}(x, s) T_{to}(x, s). \tag{4.90}$$

$$Q_{m|todp}(x, s) = \frac{\exp[\beta'_{(1)} z_{(1)}^{mtodp}]}{\sum_{m'} \exp[\beta'_{(1)} z_{(1)}^{m'todp}]}. \tag{4.91}$$

$$Q_{d|top} = \frac{\sum_{m} \exp[\beta'_{(1)} z_{(1)}^{mtodp} + \beta'_{(2)} z_{(2)}^{todp}]}{\sum_{d'} \sum_{m} \exp[\beta'_{(1)} z_{(1)}^{mtod'p} + \beta'_{(2)} z_{(2)}^{tod'p}]}, \tag{4.92}$$

and

$$Q_{p|to} = \frac{\sum_{d} \sum_{m} \exp[\beta'_{(1)} z_{(1)}^{mtodp} + \beta'_{(2)} z_{(2)}^{todp} + \beta'_{(3)} z_{(3)}^{top}]}{\sum_{p'} \sum_{d} \sum_{m} \exp[\beta'_{(1)} z_{(1)}^{mtodp'} + \beta'_{(2)} z_{(2)}^{todp'} + \beta'_{(3)} z_{(3)}^{top'}]}. \tag{4.93}$$

$$u(y, x, s) = y + \phi(x, s) + \eta(x, s), \tag{4.94}$$

$$y(p) = y_1 + (pD(p) - p_1 D_1) - C(p), \tag{4.95}$$

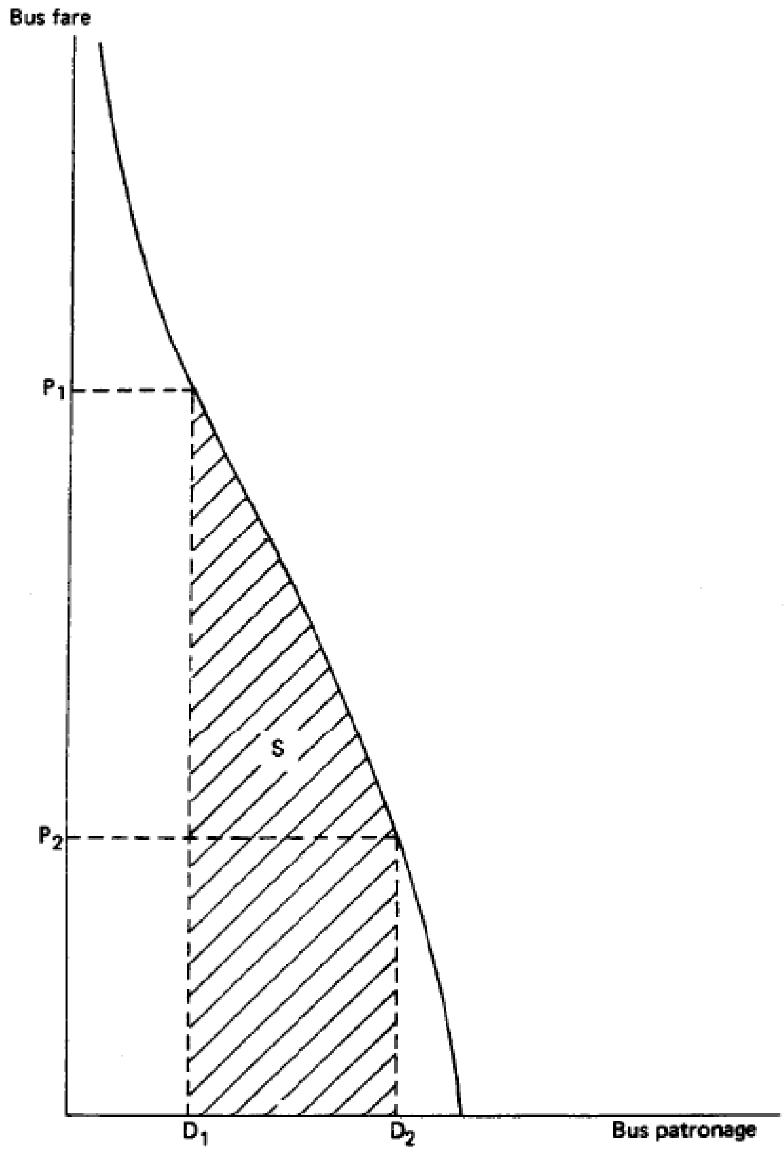


Fig. 4.6.

$$\eta(x^0, s) - \eta(x^1, s) < \phi(x^1, s) - \phi(x^0, s) - p. \tag{4.96}$$

Letting $G$ denote the cumulative distribution function of $\eta(x^0, s) - \eta(x^1, s)$, the demand for transit is

$$D(p) = G(\phi(x^1, s) - \phi(x^0, s) - p). \tag{4.97}$$

$$S = \int_{p_2}^{p_1} G(\phi(x^1, s) - \phi(x^0, s) - p)\mathrm{d}p + p_2 D_2 - p_1 D_1. \quad (4.98)$$

The utility-maximizing consumer will have a utility level

$$\begin{aligned} u(p) &= \mathrm{Max}[y(p) + \phi(x^0, s) + \eta(x^0, s), \\ &\qquad y(p) - p + \phi(x^1, s) + \eta(x^1, s)] \\ &= \phi(x^0, s) + \eta(x^1, s) + y(p) \\ &\qquad + \mathrm{Max}[\eta, \dot{\phi}(x^1, s) - \phi(x^0, s) - p] \end{aligned} \quad (4.99)$$

$$W = y(p) + \int_{-\infty}^{+\infty} \text{Max}[\eta, \phi(x^1, s) - \phi(x^0, s) - p]\, G'(\eta)\, \mathrm{d}\eta$$

$$= y(p) + \lambda G(\lambda) + \int_{\lambda}^{\infty} \eta G'(\eta)\, \mathrm{d}\eta, \qquad (4.100)$$

where $\lambda = \phi(x^1, s) - \phi(x^0, s) - p$. Under the assumptions $E\eta = 0$ and

$\int_{-\infty}^{0} G(\eta)\, \mathrm{d}\eta < +\infty$,

$$W = y(p) + \lambda G(\lambda) + E\eta - \int_{-\infty}^{\lambda} \eta G'(\eta)\mathrm{d}\eta$$
$$= y(p) + \lambda G(\lambda) - \eta G(\eta)\Big|_{-\infty}^{\lambda} + \int_{-\infty}^{\lambda} G(\eta)\mathrm{d}\eta$$
$$= y(p) + \int_{-\infty}^{\lambda} G(\eta)\mathrm{d}\eta. \qquad (4.101)$$

Considering the change in welfare resulting from a fare change from $p_1$ to $p_2$, we obtain from either eq. (4.100) or eq. (4.101) the result:

$$\Delta W = y(p_2) - y(p_1) + \int_{\lambda_1}^{\lambda_2} G(\eta)\mathrm{d}\eta, \qquad (4.102)$$

where

$$\lambda_i = \phi(x^1, s) - \phi(x^0, s) - p_i,$$

or

$$\Delta W = y_1 + (p_2 D_2 - p_1 D_1) - C - y_1$$
$$+ \int_{p_2}^{p_1} G(\phi(x^1, s) - \phi(x^0, s) - p)\mathrm{d}p = S - C. \qquad (4.103)$$

# Chapter 5
# Statistical Estimation of Choice Probability Functions

$$P_{1i} = G(V_{1i} - V_{2i}), \tag{5.1}$$

and

$$V_{ji} = V(x^{ji}, s^i) = \sum_{k=1}^{V} \beta_k Z^k(x^{ji}, s^i) = \beta' Z(x^{ji}, s^i), \tag{5.2}$$

$$P_{1i} = G(\beta' z^{1i} - \beta' z^{2i}). \tag{5.3}$$

The procedure for estimating the linear probability model is the simplest from a computational point of view. From eq. (4.11), this model can be written:

$$P_{li} = \begin{cases} 0 & \text{if } \beta' z^i < 0, \quad (5.4a) \\ \beta' z^i & \text{if } 0 \leqq \beta' z^i < 1, \quad (5.4b) \\ 1 & \text{if } 1 < \beta' z^i. \quad (5.4c) \end{cases}$$

$$f_{1i} = \beta' z^i + \varepsilon_i, \tag{5.5}$$

with $Ef_{1i} = P_{1i}$ implying $E\varepsilon_i = 0$. The estimates are then

$$\hat{\beta} = \left[ \sum_{i=1}^{I} z^i z^{i'} \right]^{-1} \left[ \sum_{i=1}^{I} z^i f_{1i} \right]. \tag{5.6}$$

$$[\hat{P}_{1i}\hat{P}_{2i}]^{-1/2} f_{1i} = [\hat{P}_{1i}\hat{P}_{2i}]^{-1/2} z^{i'} \beta, \tag{5.7}$$

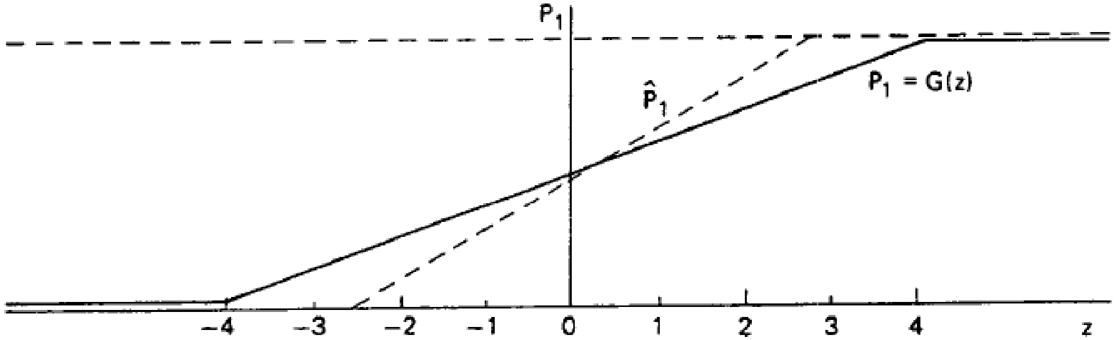



Fig. 5.1.

**True model**

$$P_{1i} = \begin{cases} 1 & \text{for} \quad z^i \geqq 4 \\ 1/2 + z^i/8 & \text{for} \quad |z^i| < 4 \\ 0 & \text{for} \quad z^i \leqq -4 \end{cases}$$

**Sample**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $z^i$ | −3 | −2 | −1 | −1 | −1 | 0 | 0 | 1 | 1 | 1 | 2 | 3 |
| $f_{1i}$ | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| $P_{1i}$ | 1/8 | 1/4 | 3/8 | 3/8 | 3/8 | 1/2 | 1/2 | 5/8 | 5/8 | 5/8 | 3/4 | 7/8 |

**Unconstrained ordinary least squares estimate**

$$\hat{P}_{1i} = \begin{cases} 1 & \text{for} \quad z^i \geqq 2\tfrac{2}{3} \\ 1/2 + 3z^i/16 & \text{for} \quad |z^i| < 2\tfrac{2}{3} \\ 0 & \text{for} \quad z^i \leqq -2\tfrac{2}{3} \end{cases}$$

The linear expression $\frac{1}{2} + 3z^i/16$ exceeds one at the data point $z^{12} = 3$.

As an alternative to this procedure, we might estimate $\beta$ in eq. (5.5) by least squares, subject to the inequality constraints

$$0 \leqq \beta' z^i \leqq 1. \tag{5.8}$$

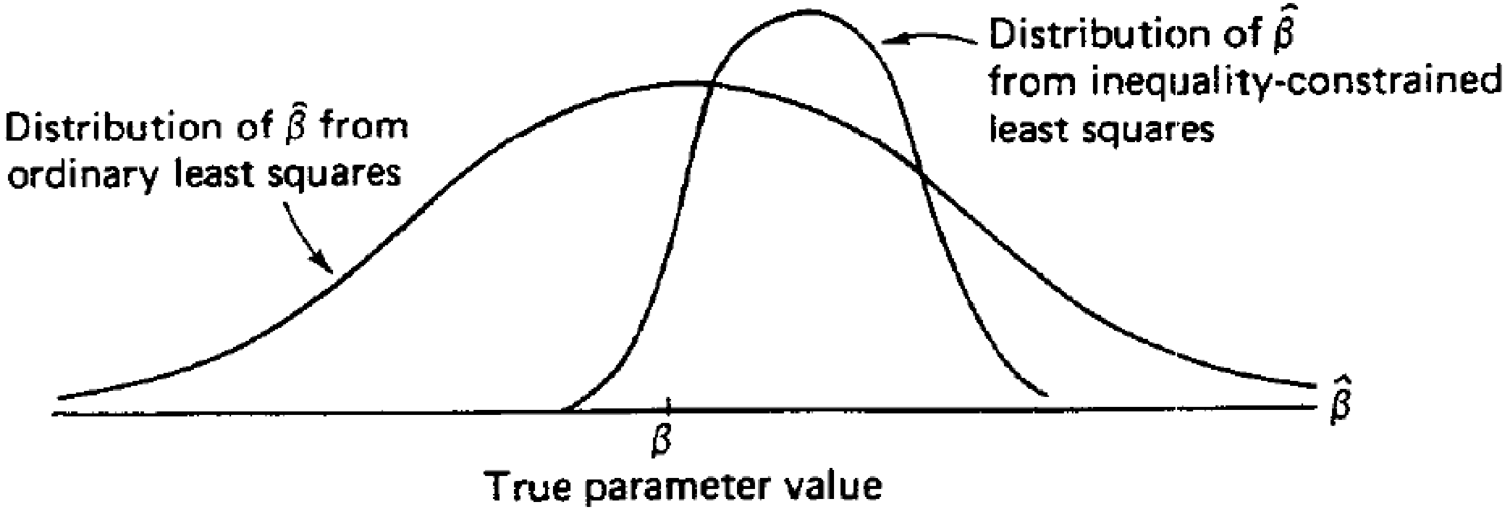



Fig. 5.2. Comparison of distribution of parameter estimates.

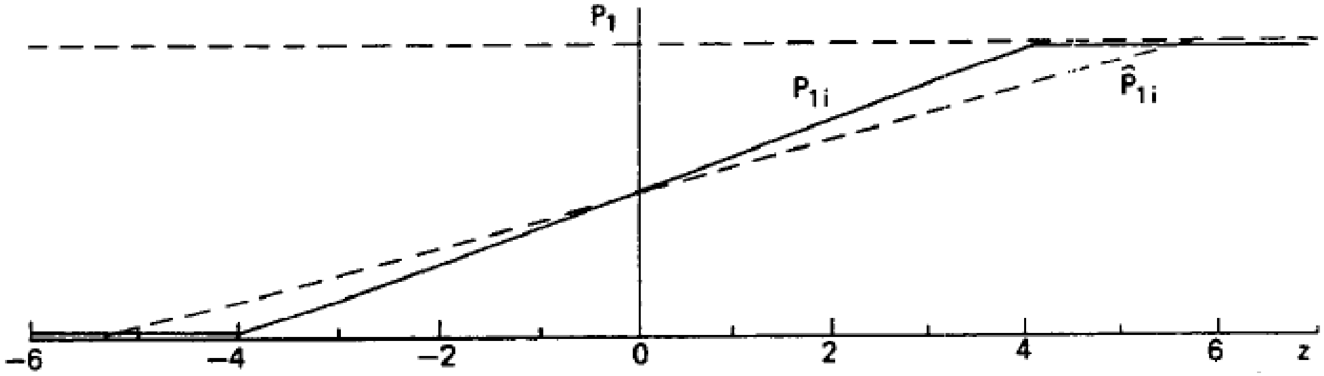



Fig. 5.3.

**True model**

$$P_{1i} = \begin{cases} 1 & \text{for} \quad z^i \geqq 4 \\ 1/2 + z^i/8 & \text{for} \quad |z^i| < 4 \\ 0 & \text{for} \quad z^i \leqq -4 \end{cases}$$

**Sample**

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $z^i$ | −7 | −5 | −3 | −2 | −1 | −1 | −1 | 0 | 0 | 1 | 1 | 1 | 2 | 3 | 5 | 7 |
| $f_{1i}$ | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |

**Ordinary least squares estimate**

$$\hat{P}_{1i} = \begin{cases} 1 & \text{for} \quad z^i \geqq 5.28 \\ 1/2 + 9z^i/95 & \text{for} \quad |z^i| < 5.28 \\ 0 & \text{for} \quad z^i \leqq 5.28 \end{cases}$$

Domencich and McFadden (1978)

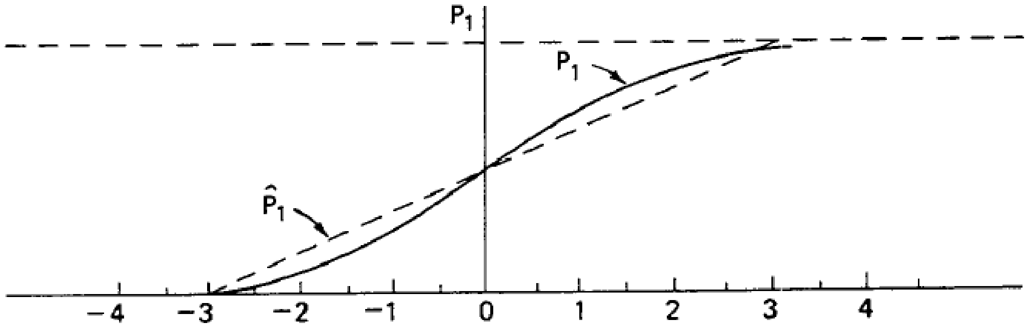



Fig. 5.4.

**True model**

$$P_{1i} = 1/(1 + e^{-z})$$

**Sample**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $i$ | 1–100 | 101–200 | 201–300 | 301–400 | 401–500 | 501–600 | 601–700 |
| $z^i$ | −3 | −2 | −1 | 0 | 1 | 2 | 3 |
| no. times 1 chosen | 5 | 12 | 27 | 50 | 73 | 88 | 95 |
| $P_{1i}$ | 0.047 | 0.119 | 0.267 | 0.5 | 0.731 | 0.881 | 0.953 |

Ordinary least squares estimation of linear probability approximation to true model.

$$\hat{P}_i = \begin{cases} 1 & \text{for } z^i \geqq 3 \\ 1/2 + 0.167\, z^i & \text{for } |z^i| < 3 \\ 0 & \text{for } z^i \leqq -3 \end{cases}$$

**Comparison of true and fitted probabilities**

| $z$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P_1$ | 0.5 | 0.731 | 0.881 | 0.953 | 0.981 |
| $\hat{P}_1$ | 0.5 | 0.67 | 0.83 | 1.0 | 1.0 |

Comparison of true and fitted effect on aggregate demand for alternative 1 of a one-unit increase in each $X_i$: true = 93, predicted = 100.

$$g(r_i/R_i) = g(P_{1i}) + g'(\rho_i)\left[\frac{r_i}{R_i} - P_{1i}\right], \tag{5.9}$$

$$g(r_i/R_i) = \beta' z^i + \varepsilon_i, \tag{5.10}$$

The case of the model (5.10) most commonly treated in the literature is the logistic distribution, yielding

$$\log\left[\frac{r_i}{R_i - r_i}\right] = \beta' z^i + \varepsilon_i. \tag{5.11}$$

Several modifications can be made in eq. (5.11) to improve the accuracy of the estimates. Cox (1970) shows that a slightly improved normal approximation and adjustment for heteroskedasticity can be attained by applying ordinary least squares to the model

$$\log\left[\frac{r_i + 1/2}{R_i - r_i + 1/2}\right] = \beta' z^i + \varepsilon_i, \tag{5.12}$$

using the resulting estimator $\bar{\beta}$ to calculate consistent estimates of the probabilities

$$P_i = 1/(1 + e^{-\bar{\beta}' z^i}),$$

and then applying least squares to the model

$$w_i \log\left[\frac{r_i + 1/2}{R_i - r_i + 1/2}\right] = \beta' z^i w_i + \varepsilon_i, \tag{5.13}$$

where

$$w_i = \sqrt{[R_i P_i (1 - P_i)]}.^2$$

$$L = \sum_{i=1}^{I} [f_{1i} \log P_{1i} + (1 - f_{1i}) \log (1 - P_{1i})]$$

$$= \sum_{i=1}^{I} \log (1 + \exp (\beta' z^i)) + \sum_{i=1}^{I} f_{1i} \beta' z^i. \qquad (5.14)$$

$$\partial L/\partial \beta = \sum_{i=1}^{I} (f_{1i} - P_{1i}) z^i = 0. \qquad (5.15)$$

The second-order condition for a maximum is

$$\frac{\partial^2 L}{\partial \beta \partial \beta'} = -\sum_{i=1}^{I} z^i P_{1i}(1 - P_{1i}) z^{i\prime} < 0. \qquad (5.16)$$

$$\frac{-E\partial^2 L}{\partial\beta\partial\beta'} = \sum_{i=1}^{I} z^i P_{1i}(1 - P_{1i})z^{i\prime}, \tag{5.17}$$

$$\log\frac{P_{1i}}{1 - P_{1i}} = \beta_1 z_1^i, \tag{5.18}$$

Table 5.1

Small sample properties of the maximum likelihood estimator (MLE), Berkson estimator (BE), and linear probability model estimator (LPE).

| Sample size | Bias | | | Mean square error | | | Berkson cell size |
|---|---|---|---|---|---|---|---|
| | MLE | BE | LPE | MLE | BE | LPE | |
| **Case A** | | | | | | | |
| 30 | 0.008 | −0.243 | −0.565 | 0.127 | 0.111 | 0.332 | 3 |
| 60 | 0.010 | −0.103 | −0.549 | 0.069 | 0.056 | 0.461 | 5 |
| 120 | −0.062 | −0.135 | −0.550 | 0.071 | 0.079 | 0.313 | 6 |
| 240 | 0.096 | −0.001 | −0.508 | 0.027 | 0.011 | 0.259 | 6 |
| **Case B** | | | | | | | |
| 30 | 0.098 | −0.319 | −0.742 | 0.181 | 0.138 | 0.551 | 10 |
| 60 | −0.050 | −0.295 | −0.742 | 0.045 | 0.100 | 0.551 | 15 |
| 120 | −0.005 | −0.276 | −0.776 | 0.034 | 0.083 | 0.603 | 20 |
| 240 | −0.062 | −0.214 | −0.774 | 0.014 | 0.052 | 0.599 | 40 |

*Note:* The independent variable has a logistic distribution with mean zero and semi-interquartile range equal to 1.10 in case A and 4.39 in case B. The true parameter value is 1.0. Twenty Monte Carlo trials are calculated for each sample size. The Berkson estimator is reported for the cell size giving the minimum mean square error.

Table 5.2

Variation of the mean square error of the Berkson estimator with changing cell size.

Case A: Semi-interquartile range of $z^i = 1.10$

| Sample size | Number of cells | Cell size | Mean square error |
|---|---|---|---|
| 30 | 30 | 1 | 0.33 |
| | 15 | 2 | 0.16 |
| | 10 | 3 | 0.11 |
| | 6 | 5 | 0.11 |
| | 5 | 6 | 0.13 |
| | 3 | 10 | 0.17 |
| 60 | 60 | 1 | 0.30 |
| | 30 | 2 | 0.13 |
| | 20 | 3 | 0.08 |
| | 15 | 4 | 0.07 |
| | 12 | 5 | 0.06 |
| | 10 | 6 | 0.08 |
| | 6 | 10 | 0.10 |
| 120 | 120 | 1 | 0.31 |
| | 60 | 2 | 0.14 |
| | 40 | 3 | 0.09 |
| | 30 | 4 | 0.07 |
| | 24 | 5 | 0.08 |
| | 20 | 6 | 0.08 |
| | 15 | 8 | 0.08 |
| | 12 | 10 | 0.08 |
| | 10 | 12 | 0.10 |
| | 6 | 20 | 0.11 |
| 240 | 240 | 1 | 0.26 |
| | 120 | 2 | 0.08 |
| | 80 | 3 | 0.03 |
| | 60 | 4 | 0.02 |
| | 48 | 5 | 0.01 |
| | 40 | 6 | 0.01 |
| | 30 | 8 | 0.01 |
| | 24 | 10 | 0.02 |
| | 20 | 12 | 0.02 |
| | 12 | 20 | 0.02 |
| | 6 | 40 | 0.02 |
| | 4 | 60 | 0.02 |

Table 5.2 (continued)

| Case B: Semi-interquartile range of $z^i = 4.39$ | | | |
|---|---|---|---|
| Sample size | Number of cells | Cell size | Mean square error |
| 30 | 30 | 1 | 0.55 |
| | 15 | 2 | 0.38 |
| | 10 | 3 | 0.28 |
| | 6 | 5 | 0.19 |
| | 5 | 6 | 0.17 |
| | 3 | 10 | 0.14 |
| 60 | 60 | 1 | 0.55 |
| | 15 | 4 | 0.22 |
| | 10 | 6 | 0.16 |
| | 6 | 10 | 0.10 |
| | 5 | 12 | 0.10 |
| | 4 | 15 | 0.11 |
| | 3 | 20 | 0.08 |
| 120 | 120 | 1 | 0.60 |
| | 30 | 4 | 0.27 |
| | 20 | 6 | 0.20 |
| | 12 | 10 | 0.13 |
| | 10 | 12 | 0.11 |
| | 8 | 15 | 0.10 |
| | 6 | 20 | 0.08 |
| | 5 | 24 | 0.09 |
| | 4 | 30 | 0.10 |
| | 3 | 40 | 0.12 |
| 240 | 240 | 1 | 0.60 |
| | 60 | 4 | 0.27 |
| | 40 | 6 | 0.19 |
| | 24 | 10 | 0.12 |
| | 20 | 12 | 0.11 |
| | 16 | 15 | 0.08 |
| | 12 | 20 | 0.07 |
| | 10 | 24 | 0.07 |
| | 8 | 30 | 0.06 |
| | 6 | 40 | 0.05 |
| | 5 | 48 | 0.07 |
| | 4 | 60 | 0.07 |
| | 3 | 80 | 0.09 |

$$\log(P_{ji}/P_{1i}) = \beta'(z^{ji} - z^{1i}), \tag{5.19}$$

or

$$P_{ji} = \frac{1}{\sum_{k=1}^{J_i} \exp(\beta'(z^{ki} - z^{ji}))}. \tag{5.20}$$

When there are repetitions at each level of the vector of explanatory variables, eq. (5.19) can be adapted to a Berkson-type analysis, with $\beta$ estimated by least squares applied to the equation

$$\log\left[\frac{r_{ji}+1/2}{r_{1i}+1/2}\right]=\beta'(z^{ji}-z^{1i})+\varepsilon_{ji}, \tag{5.21}$$

$$L=-\sum_{i=1}^{I}\sum_{j=1}^{J_i} f_{ji}\log\left[\sum_{k=1}^{J_i}\exp(\beta'(z^{ki}-z^{ji}))\right], \tag{5.22}$$

$$\partial L/\partial\beta = \sum_{i=1}^{I}\left[\sum_{j=1}^{J_i}(f_{ji} - P_{ji})z^{ji}\right], \tag{5.23}$$

and

$$\partial^2 L/\partial\beta\partial\beta' = -\sum_{i=1}^{I}\sum_{j=1}^{J_i}(z^{ji} - \bar{z}^i)P_{ji}(z^{ji} - \bar{z}^i)', \tag{5.24}$$

where

$$\bar{z}^i = \sum_{j=1}^{J_i} z^{ji}P_{ji}, \tag{5.25}$$

and

$$P_{ji} = \frac{\exp(\beta' z^{ji})}{\sum_{k=1}^{J_i}\exp(\beta' z^{ki})}. \tag{5.26}$$

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} (f_{ji} - P_{ji}) z^{ji} = 0, \tag{5.27}$$

A typical Newton–Raphson iteration, starting from a candidate parameter vector $\bar{\beta}$ and associated probabilities $\bar{P}_{ji}$ from eq. (5.26), has the form

$$\hat{\beta} = \bar{\beta} + \left[ \sum_{i=1}^{I} \sum_{j=1}^{J_i} (z^{ji} - \bar{z}^i) \bar{P}_{ji} (z^{ji} - \bar{z}^i)' \right]^{-1} \cdot \left[ \sum_{i=1}^{I} \sum_{j=1}^{J_i} (z^{ji} - \bar{z}^i)(f_{ji} - \bar{P}_{ji}) \right]. \tag{5.28}$$

Note that $\hat{\beta}$ can be interpreted as the ordinary least squares estimator in the linear model

$$\sqrt{(P_{ji})} \cdot (f_{ji} - \bar{P}_{ji}) = \sqrt{(P_{ji})} \cdot \beta'(z^{ji} - \bar{z}^i) + \varepsilon_{ji}, \tag{5.29}$$

$$S(\beta) = \sum_{i=1}^{I} \sum_{j=1}^{J_i} (f_{ji} - P_{ji}(\beta))^2 / P^*_{ji}, \tag{5.30}$$

$$R^2 = 1 - S(\hat{\beta}) / S(\overline{\beta}), \tag{5.31}$$

$$L(\beta) = \sum_{i=1}^{I} \sum_{j=1}^{J_i} f_{ji} \log P_{ji}(\beta). \tag{5.32}$$

$$\rho^2 = 1 - L(\hat{\beta}) / L(\overline{\beta}), \tag{5.33}$$

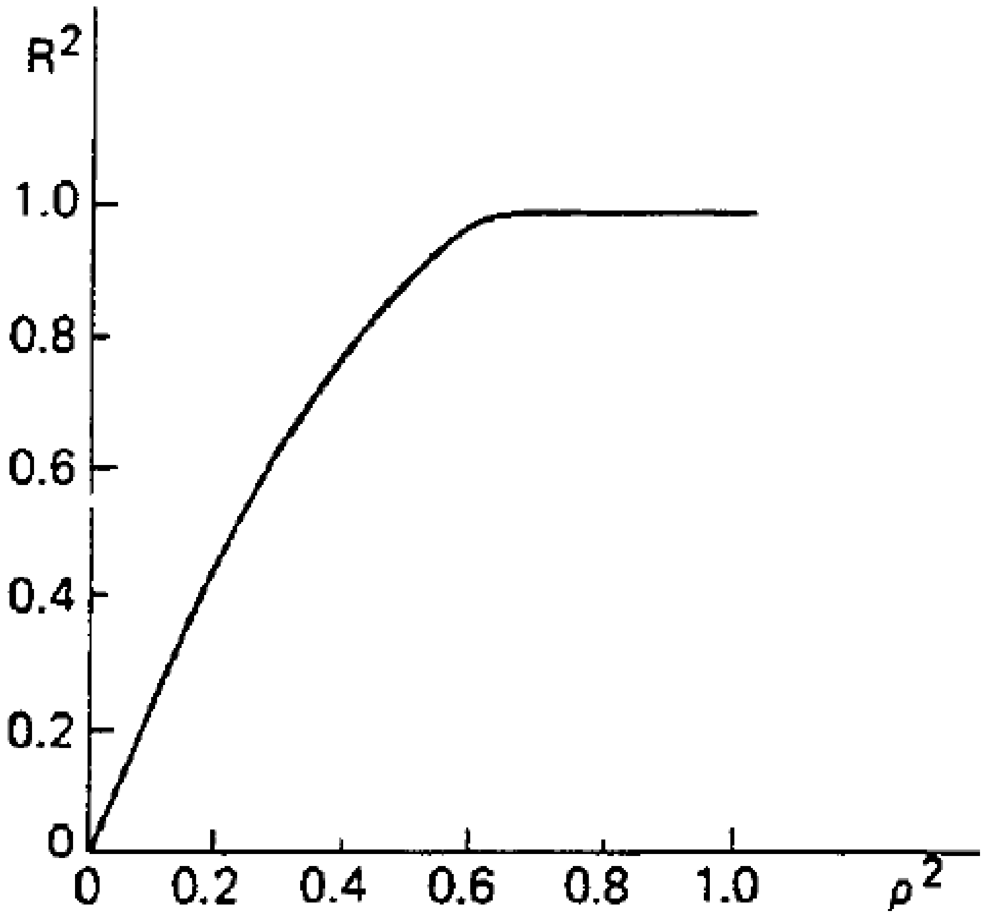



Fig. 5.5.

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J_i} f_{ji} c_j (1 - \delta_{ji}). \tag{5.34}$$

Then, the expected cost of misclassification,

$$EC = \sum_{i=1}^{I} \sum_{j=1}^{J_i} P_{ji} c_j (1 - \delta_{ji}), \tag{5.35}$$

$$\lambda = 1 - \frac{\sum_{i=1}^{I} \sum_{j=1}^{J_i} f_{ji}(1 - \delta_{ji})/\overline{P}_j}{\sum_{i=1}^{I} \sum_{j=1}^{J_i} f_{ji}(1 - 1/J_i)/\overline{P}_j}, \tag{5.36}$$