## Binary probit panel model

### Definition

Net utility from taking action in $t$ is $U_{nt} = V_{nt} + \varepsilon_{nt}$ and person takes action if $U_{nt} > 0$. This is the net utility, since it is difference between utility of taking action and that of not taking action. The errors are correlated over time, and the covariance matrix for $\varepsilon_n = (\varepsilon_{n1}, \dots, \varepsilon_{nT})$ is $\Omega \in \mathbb{R}^{T\times T}$. The sequence of actions that person $n$ takes can be represented by a set of $T$ dummy variables: $d_{nt} = 1$ if person $n$ took action at $t$ and $d_{nt} = -1$ otherwise. The probability of sequence if choices $d_n = (d_{n1}, \dots, d_{nT})$ is

$$\begin{aligned} P(d_n|V_n) &= P(U_{nt}d_{nt} > 0 \; |V_n, \; \forall t) \\ &= P(V_{nt}d_{nt} + \varepsilon_{nt}d_{nt} > 0 \; |V_n \; \forall t) \\ &= \int_{\varepsilon_n \in B_n} \phi_\Omega(\varepsilon_n) d\varepsilon_n \end{aligned}$$

where $B_n$ is a set such that $V_{nt}d_{nt} + \varepsilon_{nt}d_{nt} > 0$ **for all** $t$ and $\phi_\Omega(\varepsilon_n)$ is a joint normal density with covariance matrix $\Omega$. Once again, mathematically:

$$B_n = \{\varepsilon_n = (\varepsilon_{n1}, \dots, \varepsilon_{nT}) \mid V_{nt}d_{nt} + \varepsilon_{nt}d_{nt} > 0 \quad \forall t = 1...T\}$$

### Structure on error covariance

Structure can be placed on the covariance of the errors over time. Further we will suppose, that the error consists of a portion that is specific to the decision maker, reflecting his proclivity to take the action, and a part that varies over time for each decision maker: $\varepsilon_{nt} = \eta_n + \mu_{nt}$, where $\mu_{nt}$ is iid over time and people with a standard normal density, and $\eta_n$ is iid over people with a normal density with zero mean and variance $\sigma$. And $V(\varepsilon_{nt}) = V(\eta_n + \mu_{nt}) = V(\eta_n) + V(\mu_{nt}) = \sigma + 1$. Covariance between two time periods $s$ and $t$ is

$$\mathrm{Cov}(\varepsilon_{nt}, \varepsilon_{ns}) = \mathrm{Cov}(\eta_n + \mu_{nt}, \eta_n + \mu_{ns}) = \sigma$$

Therefore covariance matrix $\Omega$ takes form of:

$$\Omega = \begin{pmatrix} \sigma+1 & \sigma & \dots & \sigma \\ \sigma & \sigma+1 & \dots & \sigma \\ \dots & \dots & \dots & \dots \\ \sigma & \dots & \dots & \sigma+1 \end{pmatrix}$$

Only one parameter $\sigma$, enters the covariance matrix. its value indicates the variance in unobserved utility across indidividuals (the variance of $\eta_n$) relative to the variance across time for each individual (the variance of $\mu_{nt}$).

### Estimation of the parameters

In order to estimate parameters we need to simulate the choice probabilities. It is advised by *Train* to use error partioning approach. Conditional on $\eta_n$, the probability of not taking the action in period $t$ is $P(V_{nt} + \eta_n + \mu_{nt} < 0) = P(\mu_{nt} < -(V_{nt} + \eta_n)) = \Phi(-(V_{nt} + \eta_n))$, where $\Phi$ is cdf of standard normal distribution. Probability of taking action is then: $1 - \Phi(-(V_{nt} + \eta_n)) = \Phi(V_{nt} + \eta_n)$. The probability of the sequence of choices $d_n = (d_{n1}, \ldots, d_{nT})$ is then:

$$P(d_n|\eta_n, V_n) = \prod_{t=1}^{T} \Phi((V_{nt} + \eta_n)d_{nt})$$

So far we have conditioned on $\eta_n$, when in fact $\eta_n$ is random. The unconditional probability is the integral of the conditional probability $P(d_n|\eta_n, V_n)$ over all possible values of $\eta_n$:

$$P(d_n|V_n) = \int_{\mathbb{R}} P(d_n|\eta_n, V_n)\phi_\sigma(\eta_n)d\eta_n$$

where $\phi_\sigma(\eta_n)$ is normal density with zero mean and variance $\sigma$. Overall result:

$$P(d_n|V_n) = \int_{\mathbb{R}} \prod_{t=1}^{T} \Phi((V_{nt} + \eta_n)d_{nt})\phi_\sigma(\eta_n)d\eta_n$$

Once again $P(d_n|V_n)$ is unconditional (w.r.t $\eta_n$) probability that agent $n$ will choose sequence of choices $d_n = (d_{n1}, \ldots, d_{nT})$, where each action $d_{nt}$ is either 1 if action is taken and $-1$ otherwise. *Train* (ch 5, pg. 17) proposes to use the following algorithm for $P(d_n|V_n)$ estimation:

1. Take a draw of $\eta_n$ from $\phi_\sigma(\eta_n)$.
2. For this draw of $\eta_n$, calculate $P(d_n|\eta_n, V_n)$ .
3. Repeat steps 1–2 many times, and average the results. This average is a simulated approximation to $P(d_n|V_n)$ .