- *ordered probit* model (F is normal)
  - ➢ If the latent variable $z_{\mathbf{x}}$ has a standard normal distribution, then

$$\Phi^{-1}(\gamma_j(\mathbf{x})) = \beta_{0j} - \boldsymbol{\beta}^T\mathbf{x}^*$$

on cdf, not pmf

  - ➢ Coefficients estimates could be very different from proportional odds model, but predicted values usually are very similar $P_{x.1}, P_{x.J} \geq 0.2$
- *proportional hazards* model (F = extreme value dist. hazard = $\exp(z = t_j - \beta^T x^*)$) [$\Phi(z) \propto e^z/(1+e^z)$, $0.2 \leq \Phi(z) \leq 0.8$]
  - ➢ Concept of hazard [pdf ⇔ survival function ⇔ cdf ⇔ hazard]

proportion hazard: $h_x(t) = h_0(t) \cdot \exp(-\beta^T x^*)$ ⇔ ⓑ $S_x(t) = S_0(t)^{\exp(-\beta^T x^*)}$ ← survival function at x

    - Developed in insurance application: when issuing a life insurance policy, the insurer is interested in the prob. that the person will die during the term of the policy given that they are alive now ← $P(\text{die at } 55 \mid \text{age} \geq 55)$ ↔ hazard function ⓓ $\frac{S_x(t_2)}{S_x(t_1)} = \left[\frac{S_0(t_2)}{S_0(t_1)}\right]^{\exp(-\beta^T x^*)}$
    - This is not the same as the unconditional probability of death ← pdf ⟷ $P(\text{die at } 55)$

[Consider the case z_x (LN p.6-7): lifetime & F: a lifetime distribution]

  - ➢ Suppose we use the complementary log-log link, i.e.,

ⓐ $\frac{h_{x_1}(t)}{h_{x_2}(t)} = \exp[-\beta^T(x_1^* - x_2^*)]$ irrelevant to t.

$\log(1-\gamma_{x.j}) = -\exp(\beta_{0j} - \beta^T x^*)$ ⇐
$$\log(-\log(1-\gamma_j(\mathbf{x}))) = \beta_{0j} - \boldsymbol{\beta}^T\mathbf{x}^*$$

For x_0 s.t. $x^* = 0$: $\log(1-\gamma_{x_0.j}) = -\exp(\beta_{0j})$ ⇓

Then, the hazard of category j is the probability of falling in category j given that your category is greater than j: ⓒ $\frac{\log(S_{x_1}(t))}{\log(S_{x_2}(t))} = \exp[-\beta^T(x_1^* - x_2^*)]$

$$\text{hazard}_{\mathbf{x}}(j) = P(y_{\mathbf{x}} = j \mid y_{\mathbf{x}} \geq j) = P(y_{\mathbf{x}} = j)/P(y_{\mathbf{x}} \geq j)$$

$$= \frac{p_{\mathbf{x}j}}{1-\gamma_{\mathbf{x},j-1}} = \frac{\gamma_{\mathbf{x}j} - \gamma_{\mathbf{x},j-1}}{1-\gamma_{\mathbf{x},j-1}} = 1 - \frac{1-\gamma_{\mathbf{x}j}}{1-\gamma_{\mathbf{x},j-1}}$$

$\log(1-\gamma_{x.j}) = \log(1-\gamma_{x_0.j}) \cdot \exp(-\beta^T x^*)$ ⇒ $1-\gamma_{x.j} = (1-\gamma_{x_0.j})^{\exp(-\beta^T x^*)}$ ⓑ

Say. $x \to x+1$

ⓐ not hold

  - These hazards are then proportional across categories as **x** varies
  - The corresponding latent variable distribution is the extreme value distribution: ⓓ $\frac{1-\gamma_{x.j}}{1-\gamma_{x.j-1}} = \frac{P(y_x \geq j+1)}{P(y_x \geq j)} = \left(\frac{1-\gamma_{x_0.j}}{1-\gamma_{x_0.j-1}}\right)^{\exp(-\beta^T x^*)}$

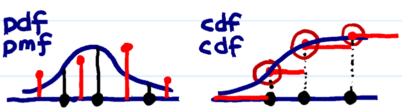
$$F(z) = 1 - \exp(-\exp(z))$$

check its graph in LN p.6-7 ← ; $\log(-\log(1-h_{x.j})) = \beta_{0j}^* - \beta^T x^*$

[$\frac{\log(1-\gamma_{x_1.j})}{\log(1-\gamma_{x_2.j})} = \exp[-\beta^T(x_1^* - x_2^*)]$ irrelevant to j. ⓒ]

  - The extreme value distribution is not symmetric like the logistic and normal

[$F(z) \uparrow 1$ as $z \to \infty$; $F(z) \downarrow 0$ as $z \to -\infty$] faster



- Generalization:
  - ➢ The proportional hazards and odds models can be generalized by allowing beta to vary, i.e.,

[Can be used as the alternative model to test whether $\beta_1 = \cdots = \beta_J$] →
$$\log\left(\frac{\gamma_j(\mathbf{x})}{1-\gamma_j(\mathbf{x})}\right) = \beta_{0j} - \boldsymbol{\beta}_j^T\mathbf{x}^*$$

depends on j. ; ↳ has same # of parameters as multinomial logit model.

  - ➢ But, this loses the proportionality property

❖ **Reading**: Faraway (2006, 1st ed.), 5.3

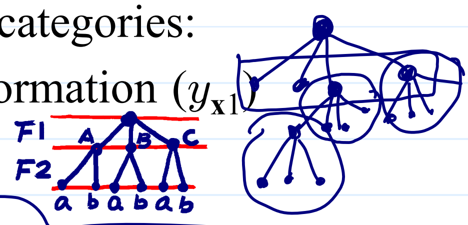
crossed Hierarchical or Nested Multinomial Responses

- Consider a multinomial response with the 4 categories:
  - **NoCNS**: no central nervous system malformation (y_{x1})

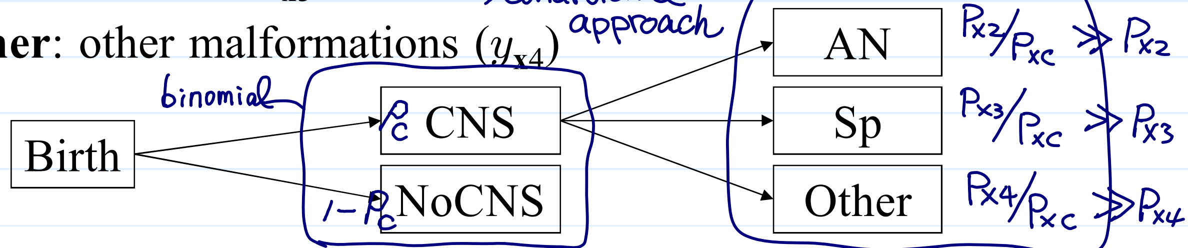
    AN: anencephalus (y_{x2})

    Sp: spina bifida (y_{x3})

    Other: other malformations (y_{x4})



multinomial ⇒ conditional approach



  - There exists a hierarchical structure between the 4 categories
  - **Q**: what are the problems if we ignore the hierarchical structure and just treat them as 4 nominal categories in the analysis?
    - In the data, most births suffer no malformation and so NoCNS dominates the other 3 categories (**Q**: why is this a problem?)
    - **Q**: what happen if **x** has significant effects on the prob. of NoCNS but not on the other 3 categories, and NoCNS is chosen to be the baseline? (check lab)

- The likelihood of (y_{x1}, y_{x2}, y_{x3}, y_{x4}) is proportional to (note: most inference methods in GLM are based on the likelihood approach):

$$p_{x1}^{y_{x1}} p_{x2}^{y_{x2}} p_{x3}^{y_{x3}} p_{x4}^{y_{x4}}$$

$$= \underbrace{\left[p_{x1}^{y_{x1}} (p_{xc})^{\overbrace{y_{x2}+y_{x3}+y_{x4}}^{m_x}} \right]}_{\text{1st part}} \times \underbrace{\left[\left(\frac{p_{x2}}{p_{xc}}\right)^{y_{x2}} \left(\frac{p_{x3}}{p_{xc}}\right)^{y_{x3}} \left(\frac{p_{x4}}{p_{xc}}\right)^{y_{x4}} \right]}_{\text{2nd part}}$$

$\oplus = 1$.

where $p_{xc}=p_{x2}+p_{x3}+p_{x4}=1-p_{x1}$ (= prob. of a birth with CNS malform.)

  - The 1st part is a binomial likelihood for CNS vs. NoCNS
  - The 2nd part is a multinomial likelihood for the three CNS categories *conditional* on the presence of CNS $y_{x2}+y_{x3}+y_{x4}$ treated as fixed number of trials
    - e.g., p_{x2}/p_{xc} is the conditional probability of an anencephalus birth given that a malformation has occurred at **x**
  - We can separately
    - develop a binomial model for whether malformation occurs using data from *all objects*, and
    - develop a multinomial model for the type of malformation using data only from *subjects with CNS malformation*

no individual information about P_{x2}, P_{x3}, P_{x4}

$Y_{x,1} \sim$ binomial$(n_x, P_{x,1})$

$(Y_{x,2}, Y_{x,3}, Y_{x,4}) \sim$ multinomial$\left(m_x, \frac{P_{x,2}}{P_{x,c}}, \frac{P_{x,3}}{P_{x,c}}, \frac{P_{x,4}}{P_{x,c}}\right)$

❖ **Reading**: Faraway (2006, 1st ed.), 5.2