SEM 2: Structural Equation Modeling

Week 3 - Partial expectation

Sacha Epskamp

If the SEM model fits, then all implied conditional independence relations are likely to hold. We can now investigate the predictive effect of seeing $\boldsymbol{x} = \boldsymbol{z}$ (some observed value) on $\boldsymbol{y}$:

$$\text{Predictive effect} = \mathcal{E}(\boldsymbol{y} \mid \text{See}(\boldsymbol{x} = \boldsymbol{z})) - \mathcal{E}(\boldsymbol{y})$$

as well as the causal effect of $\boldsymbol{x}$ on $\boldsymbol{y}$ ($\boldsymbol{z}$ is the result of a causal intervention):

$$\text{Causal effect} = \mathcal{E}(\boldsymbol{y} \mid \text{Do}(\boldsymbol{x} = \boldsymbol{z})) - \mathcal{E}(\boldsymbol{y})$$

If the SEM model fits, then all implied conditional independence relations are likely to hold. We can now investigate the predictive effect of seeing $\boldsymbol{x} = \boldsymbol{z}$ (some observed value) on $\boldsymbol{y}$:

$$\text{Predictive effect} = \mathcal{E}(\boldsymbol{y} \mid \text{See}(\boldsymbol{x} = \boldsymbol{z})) - \mathcal{E}(\boldsymbol{y})$$

as well as the causal effect of $\boldsymbol{x}$ on $\boldsymbol{y}$ ($\boldsymbol{z}$ is the result of a causal intervention):

$$\text{Causal effect} = \mathcal{E}(\boldsymbol{y} \mid \text{Do}(\boldsymbol{x} = \boldsymbol{z})) - \mathcal{E}(\boldsymbol{y})$$

For simplicity, we will assume all variables to be centered:

$$\mathcal{E}(\boldsymbol{x}) = \boldsymbol{0} \quad \& \quad \mathcal{E}(\boldsymbol{y}) = \boldsymbol{0}$$

which simply makes the predictive effect $\mathcal{E}(\boldsymbol{y} \mid \text{See}(\boldsymbol{x} = \boldsymbol{z}))$ and the causal effect $\mathcal{E}(\boldsymbol{y} \mid \text{Do}(\boldsymbol{x} = \boldsymbol{z}))$

The predictive effect can then be obtained from the conditional Gaussian distribution (assuming centered variables):

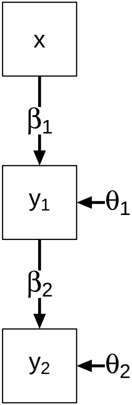
$$\mathcal{E}(\boldsymbol{y} \mid \mathrm{See}(\boldsymbol{x} = \boldsymbol{z})) = \mathrm{Cov}(\boldsymbol{y}, \boldsymbol{x})\mathrm{Var}(\boldsymbol{x})^{-1}\boldsymbol{z}$$

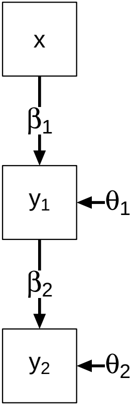
Or if x and y are both single variables:

$$\mathcal{E}(y \mid \mathrm{See}(x = x_i)) = \frac{\mathrm{Cov}(y, x)}{\mathrm{Var}(x)} \times z$$

Predictive effect of x on y_2:

$$\begin{aligned}\mathcal{E}(y_2 \mid \mathrm{See}(x = z)) &= \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(x)} \times z \\ &= \frac{\beta_2 \beta_1 \mathrm{Var}(x)}{\mathrm{Var}(x)} \times z \\ &= \beta_2 \beta_1 z\end{aligned}$$



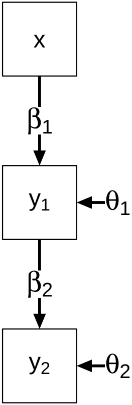


Predictive effect of x on y_2:

$$\begin{aligned}\mathcal{E}(y_2 \mid \mathrm{See}(x = z)) &= \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(x)} \times z \\ &= \frac{\beta_2 \beta_1 \mathrm{Var}(x)}{\mathrm{Var}(x)} \times z \\ &= \beta_2 \beta_1 z\end{aligned}$$

Predictive effect of y_1 on y_2:

$$\begin{aligned}\mathcal{E}(y_2 \mid \mathrm{See}(y_1 = z)) &= \frac{\mathrm{Cov}(y_2, y_1)}{\mathrm{Var}(y_1)} \times z \\ &= \frac{\beta_2 \beta_1^2 \mathrm{Var}(x) + \beta_2 \theta_1}{\beta_1^2 \mathrm{Var}(x) + \theta_1} \times z \\ &= \beta_2 z\end{aligned}$$



Predictive effect of x on y_2:

$$\mathcal{E}(y_2 \mid \mathrm{See}(x = z)) = \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(x)} \times z = \frac{\beta_2\beta_1\mathrm{Var}(x)}{\mathrm{Var}(x)} \times z = \beta_2\beta_1 z$$

Predictive effect of y_1 on y_2:

$$\mathcal{E}(y_2 \mid \mathrm{See}(y_1 = z)) = \frac{\mathrm{Cov}(y_2, y_1)}{\mathrm{Var}(y_1)} \times z = \frac{\beta_2\beta_1^2\mathrm{Var}(x) + \beta_2\theta_1}{\beta_1^2\mathrm{Var}(x) + \theta_1} \times z = \beta_2 z$$

Predictive effect of y_2 on x:

$$\mathcal{E}(x \mid \mathrm{See}(y_2 = z)) = \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(y_2)} \times z = \frac{\beta_2\beta_1\mathrm{Var}(x)}{\beta_2^2\beta_1^2\mathrm{Var}(x) + \beta_2^2\theta_1 + \theta_2} \times z$$

For the causal effect:

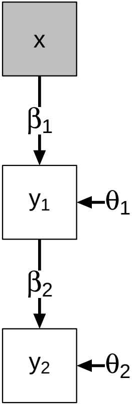
1. Make a dummy model in which the variable you intervene on is *exogenous*
   - Remove all incoming uni-directional arrows (causal effects) to the variable you intervene on
   - Remove all covariances connected to the variable you intervene on
   - Retain the *variance* of the variable you intervene on
2. Now, compute the predictive effect using this dummy model

For the causal effect:

1. Make a dummy model in which the variable you intervene on is *exogenous*
   - Remove all incoming uni-directional arrows (causal effects) to the variable you intervene on
   - Remove all covariances connected to the variable you intervene on
   - Retain the *variance* of the variable you intervene on
2. Now, compute the predictive effect using this dummy model

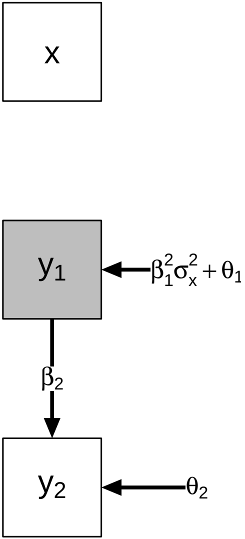
An alternative is to calculate the total effect:

1. List all paths of uni-directional edges from the node you intervene on to the node of interest (only going forward along the direction)
2. For each path: multiply all regression coefficients on the path
3. Sum all these products
4. Multiply the result with the result of your intervention



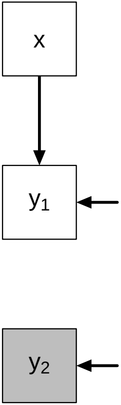
Causal effect of x on y_2 (no difference in dummy model):

$$\begin{aligned}\mathcal{E}(y_2 \mid \mathrm{Do}(x = z)) &= \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(x)} \times z \\ &= \frac{\beta_2 \beta_1 \mathrm{Var}(x)}{\mathrm{Var}(x)} \times z \\ &= \beta_2 \beta_1 z\end{aligned}$$

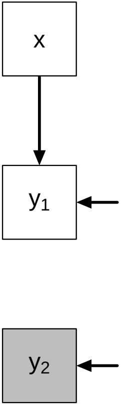


Causal effect of y_1 on y_2 (using dummy model):

$$\begin{aligned}\mathcal{E}(y_2 \mid \mathrm{Do}(y_1 = z)) &= \frac{\mathrm{Cov}(y_2, y_1)}{\mathrm{Var}(y_1)} \times z \\ &= \frac{\beta_2(\beta_1^2 \mathrm{Var}(x) + \theta_1)}{\beta_1^2 \mathrm{Var}(x) + \theta_1} \times z \\ &= \beta_2 z\end{aligned}$$



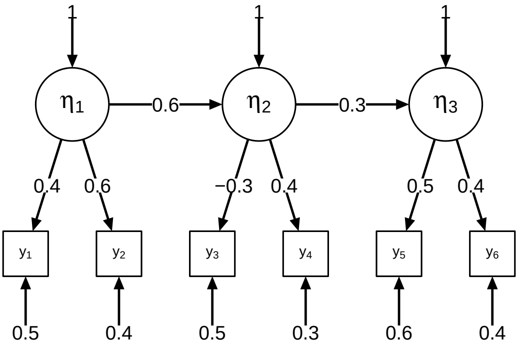
Causal effect of y_2 on x?



Causal effect of y_2 on x? Using the dummy model:

$$\mathcal{E}(x \mid \mathrm{Do}(y_2 = z)) = \frac{\mathrm{Cov}(y_2, x)}{\mathrm{Var}(y_2)} \times z$$

But now $\mathrm{Cov}(y_2, x) = 0$, and hence $\mathcal{E}(x \mid \mathrm{Do}(y_2 = z)) = 0$!



$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\begin{aligned} \mathrm{Var}(\boldsymbol{y}) = \boldsymbol{\Sigma} &= \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{B})^{-1}\boldsymbol{\Psi}(\boldsymbol{I} - \boldsymbol{B})^{-1\top}\boldsymbol{\Lambda}^\top + \boldsymbol{\Theta} \\ &= \begin{bmatrix} 0.66 & 0.24 & -0.07 & 0.10 & 0.04 & 0.03 \\ 0.24 & 0.76 & -0.11 & 0.14 & 0.05 & 0.04 \\ -0.07 & -0.11 & 0.62 & -0.16 & -0.06 & -0.05 \\ 0.10 & 0.14 & -0.16 & 0.52 & 0.08 & 0.07 \\ 0.04 & 0.05 & -0.06 & 0.08 & 0.88 & 0.22 \\ 0.03 & 0.04 & -0.05 & 0.07 & 0.22 & 0.58 \end{bmatrix} \end{aligned}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\begin{aligned} \mathrm{Var}(\boldsymbol{\eta}) &= (\boldsymbol{I} - \boldsymbol{B})^{-1} \boldsymbol{\Psi} (\boldsymbol{I} - \boldsymbol{B})^{-1\top} \\ &= \begin{bmatrix} 1.00 & 0.60 & 0.18 \\ 0.60 & 1.36 & 0.41 \\ 0.18 & 0.41 & 1.12 \end{bmatrix} \end{aligned}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\begin{aligned} \mathrm{Var}(\boldsymbol{\eta}) &= (\boldsymbol{I} - \boldsymbol{B})^{-1} \boldsymbol{\Psi} (\boldsymbol{I} - \boldsymbol{B})^{-1\top} \\ &= \begin{bmatrix} 1.00 & 0.60 & 0.18 \\ 0.60 & 1.36 & 0.41 \\ 0.18 & 0.41 & 1.12 \end{bmatrix} \end{aligned}$$

Note: η_1 is conditionally independent from η_3 given η_2:

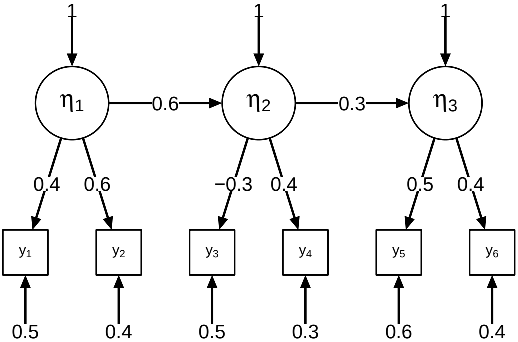
$$\begin{aligned} \mathrm{Cov}(\eta_1, \eta_3 \mid \eta_2) &= \mathrm{Cov}(\eta_1, \eta_3) - \frac{\mathrm{Cov}(\eta_1, \eta_2)\mathrm{Cov}(\eta_3, \eta_2)}{\mathrm{Var}(\eta_2)} \\ &= 0.18 - \frac{0.6 \times 0.41}{1.36} = 0 \end{aligned}$$

(rounded to two digits).

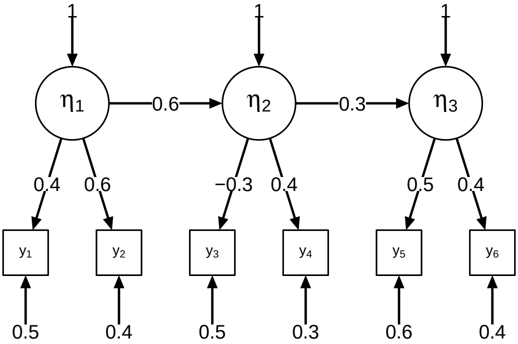
$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.4 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & -0.3 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \end{bmatrix}, \boldsymbol{\Psi} = \boldsymbol{I}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}, \mathrm{Diag}(\boldsymbol{\Theta}) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 0.3 \\ 0.6 \\ 0.4 \end{bmatrix}$$

$$\begin{aligned} \mathrm{Cov}(\boldsymbol{\eta}, \boldsymbol{y}) &= \mathrm{Cov}(\boldsymbol{\eta}, \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}) \\ &= \mathrm{Var}(\boldsymbol{\eta})\boldsymbol{\Lambda}^\top \\ &= (\boldsymbol{I} - \boldsymbol{B})^{-1}\boldsymbol{\Psi}(\boldsymbol{I} - \boldsymbol{B})^{-1\top}\boldsymbol{\Lambda}^\top \\ &= \begin{bmatrix} 0.40 & 0.60 & -0.18 & 0.24 & 0.09 & 0.07 \\ 0.24 & 0.36 & -0.41 & 0.54 & 0.20 & 0.16 \\ 0.07 & 0.11 & -0.12 & 0.16 & 0.56 & 0.45 \end{bmatrix} \end{aligned}$$



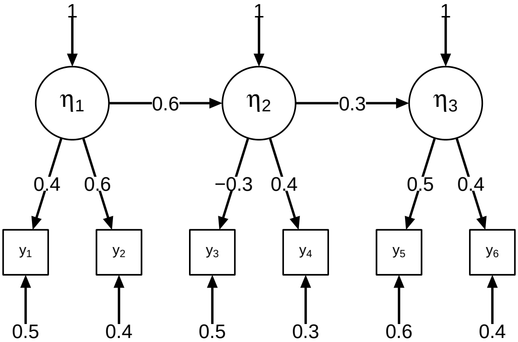
What is $\mathrm{Cov}(y_1, y_6 \mid \eta_2)$?



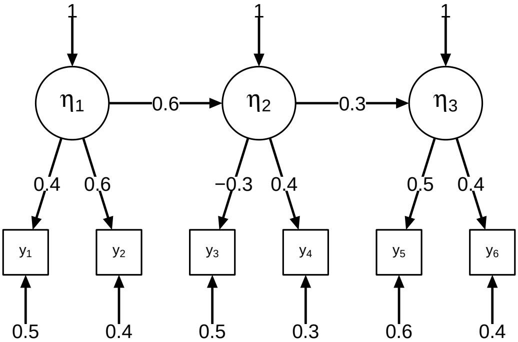
What is $\mathrm{Cov}(y_1, y_6 \mid \eta_2)$?

$$\mathrm{Cov}(y_1, y_6 \mid \eta_2) = \mathrm{Cov}(y_1, y_6) - \frac{\mathrm{Cov}(y_1, \eta_2) Cov(\eta_2, y_6)}{\mathrm{Var}(\eta_2)}$$

$$= 0.03 - \frac{0.24 \times 0.16}{1.36} = 0$$

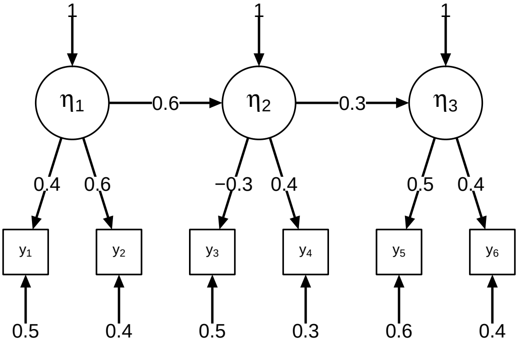


What is $\mathcal{E}(y_6 \mid \mathrm{See}(y_1 = 0.5))$?

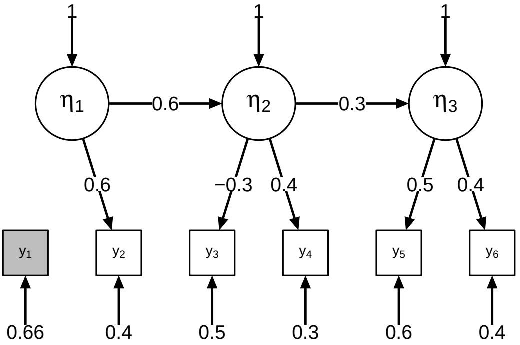


What is $\mathcal{E}(y_6 \mid \mathrm{See}(y_1 = 0.5))$?

$$\mathcal{E}(y_6 \mid \mathrm{See}(y_1 = 0.5)) = \frac{\mathrm{Cov}(y_1, y_6)}{\mathrm{Var}(y_1)} \times 0.5$$
$$= \frac{0.03}{0.66} \times 0.5 = 0.02$$

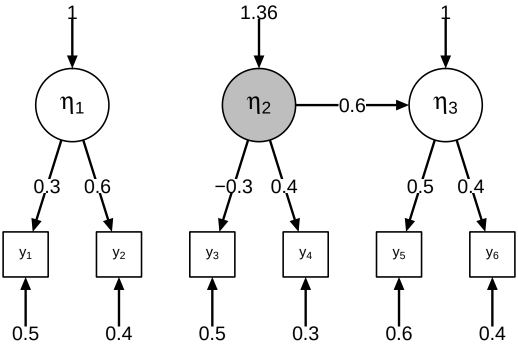


What is $\mathcal{E}(y_6 \mid \mathrm{Do}(y_1 = 0.5))$?

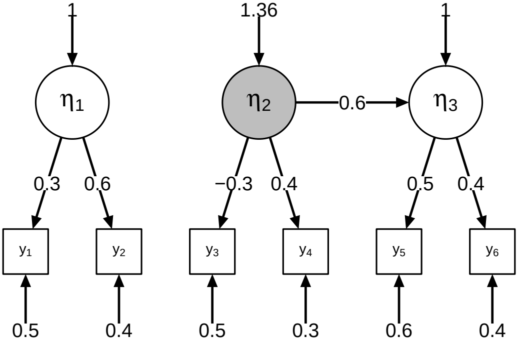


What is $\mathcal{E}(y_6 \mid \mathrm{Do}(y_1 = 0.5))$?

$$\mathcal{E}(y_6 \mid \mathrm{Do}(y_1 = 0.5)) = 0$$



What is $\mathcal{E}(y_6 \mid \mathrm{Do}(\eta_2 = 0.5))$?



What is $\mathcal{E}(y_6 \mid \mathrm{Do}(\eta_2 = 0.5))$?

$$\mathcal{E}(y_6 \mid \mathrm{Do}(y_1 = 0.5)) = \frac{0.4 \times 0.6 \times 1.36}{1.36} \times 0.5 = 0.12$$