2016 Q1

- $y_i \sim Pois(\lambda_i)$, $\lambda_i = x_i'\beta$

$$P(y_i = y) = \frac{\lambda^y \exp(-\lambda)}{y!}, \quad y \geq 0.$$

$$\Rightarrow E[y] = Var(y) = \lambda$$

1.1) $y_i = x_i'\beta + \varepsilon_i$ (vector notation)

Just the OLS estimator

$$\hat{\beta} = \mathbb{E}_n[x_i x_i']^{-1} \mathbb{E}_n[x_i y_i]$$

which we know is consistent if $E[x_i x_i']$ has full rank and

$$\begin{aligned} E[\varepsilon_i | x_i] &= E[y_i - x_i'\beta | x_i] \\ &= E[y_i | x_i] - E[x_i'\beta | x_i] \\ &= 0. \end{aligned}$$

12) $\hat{\beta}$ unbiased?

- $\hat{\beta}$ unbiased by property of OLS (since $E[\varepsilon_i | x_i] = 0$).
- $\hat{\beta}$ linear estimator as req. by Gauss-Markov.
- Not BLUE because our problem has

$$E[y_i|X_i] = \underline{Var[y_i|X_i] = \lambda_i}$$

Recall, to use G-M Thm:

1) $\hat{\beta}$ unbiased.
   $\hookrightarrow E[\hat{\beta}|X] = \beta$
2) $var(\hat{\beta}|X) = \sigma^2 (X'X)^{-1}$

However,

$$\underline{\lambda_i = var(y_i | X_i)} = var(x_i'\beta + \varepsilon_i | X_i)$$
$$= var(\varepsilon_i | x_i)$$
$$= \lambda_i$$
$$= x_i'\beta$$

1.3) Assuming $\lambda_i = \exp(x_i'\beta)$ true

$\hookrightarrow \log y_i = x_i'\beta + \varepsilon_i$ can we do this?

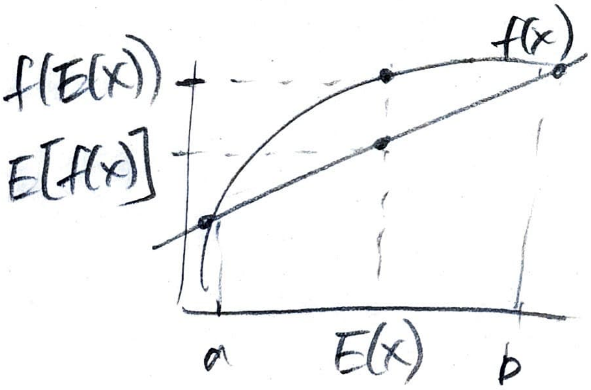
Before:

If $\lambda_i = \exp(x_i'\beta)$, then

$E[y_i | x_i] = \lambda_i = \exp(x_i'\beta)$ "true".

After.

$$E[\log y_i | x_i] \leq \log(E(y_i | x_i))$$
$$= \log(\exp(x_i'\beta))$$
$$= x_i'\beta$$



Not true by Jensen's Inequality

1.4) Estimate

$$y_i = x_i'\beta + \varepsilon \quad , \text{ when } y_i \sim Pois(\lambda_i)$$

$$\lambda_i = \exp(x_i'\beta)$$

Goal: $\hat{\beta}$

$$\lambda_i = \exp(x_i'\beta)$$

Recall, $E[y_i] = Var(y_i) = \lambda_i = \exp(x_i'\beta)$.

We can perform a nonlinear GMM estimation with moment condition.

$$E[y_i - \underbrace{\exp(x_i'\beta)}_{E(y_i) = \lambda_i}] = 0$$

2024 Q2

- assume $E[\varepsilon^2 | X] = \sigma^2 I_n$

2.1) 1) $X = \delta Z + \varepsilon$

2) $Y = \hat{\beta}\hat{X}$

$\Rightarrow$ The $\hat{\beta}$ is $\hat{\beta}_{TSLS}$ (or $\hat{\beta}_{IV}$ if $l = k$)

Requires:

- homoscedasticity
- $l \geq k$
- $E[Z\varepsilon] = 0$

2.2) 1) $\hat{X} = \hat{\delta} Z$

1.5) $X = \hat{X} + \hat{\eta}$

2) $Y = \beta_2(\hat{\eta}) + \varepsilon$, $\hat{\eta} = X - \hat{X}$

$\hat{\eta}$ = amount of variation $\hat{X}$ cannot explain in X

Stage (2) is equivalent to $Y = \beta_0 + \underbrace{\phi_1 \hat{X}}_{= \underbrace{\phi_1 \hat{\delta}}_{} Z = \beta_1 Z} + \beta_2 X + e$

Hence, this is just a typical multivariate OLS model where we require

- $l = k$
- $E[X\varepsilon] = E[Z\varepsilon] = 0$

Note: This is a Full depiction of OLS.

2.3) Obviously, if the assumptions hold for both models, either estimator works.

If we have that $E[Z\varepsilon] = 0$, but $E[X\varepsilon] \neq 0$
$\Rightarrow$ IV.

Hovever, if our goal is to simply find the causal effect of X on Y (assuming all assumptions hold), we could perform simple OLS.

Side Note: Could perform SUR if $E[Z\varepsilon] = 0$ and $l = k$.

2014 Q3

3.1) $$\hat{\theta}(\hat{W}) = (S'\hat{W}S)^{-1}S'\hat{W}s.$$

$$\Rightarrow \hat{\beta}(\hat{W}) = (X'Z\hat{W}Z'X)^{-1}(X'Z\hat{W}Z'Y).$$

$S = Z'X$

$s = Z'Y$

Note: There doesn't seem to be instruments here, so $Z = X$

$$\theta = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \qquad S = \mathbb{E}_n[X'X].$$

$$s = \mathbb{E}_n[X'Y] = \mathbb{E}_n \begin{bmatrix} x_i \varepsilon_i \\ x_i \eta_i \\ x_i v_i \end{bmatrix}$$

For $\hat{W} \xrightarrow{P} W$, where W is the efficient weighting matrix (i.e. true inverse of var cov), then my $\hat{W}$ must at least be symmetric and positive definite.

3.2) Here, we have a just-identified model ($l = k$), so the choice of $\hat{W}$ does not matter because things will cancel out.

Hence, estimating this model eq-by-eq vs. jointly are identical (assuming W is full rank).

3.3) WLOG, if $\alpha' = 0$, then we have an overidentified model in which

$$l = 3 > k = 2.$$

Hence, we should estimate jointly to attain better efficiency.

However, the drawback of estimating jointly is that misspecification will permeate throughout the model, so other estimators will be affected as well.