Extension of the corrected score estimator in a Poisson regression model with a measurement error

Kentarou Wada and Takeshi Kurosawa

Department of Applied Mathematics, Tokyo University of Science

November 26, 2025

Error-in-variables

Error-in-variables (EIV)

A regression model with measurement errors in explanatory variables.

In this research, we focus on estimation methods of a Poisson regression model with EIV.

An example of a Poisson regression with EIV

Real data examples include epidemiologic studies (e.g., [1]).

Y	:	Cardiovascular disease emergency visits
X	:	True exposure amount (air pollutants)

Consider regressing Y on X, but X is subject to measurement error U due to instrument imprecision and spatial variability.

Research status of a Poisson regression model with EIV

Table: Bias correction using estimating equations of EIV

Explanatory variable	Normal	General
Univariate	Kukush et al. [2]	Wada et al. [3]
Multivariate(All error)	Shklyar et al. [4]	Wada et al. [5]
Multivariate(Partial error)	Wada et al. [5]	Wada et al. [5]

All error : All explanatory variables are subject to errors
Partial error : Explanatory variables are partially subject to errors

We call the estimator in Wada et al. [5] the CPN estimator.

Problems of the CPN estimator

Although the CPN estimator is consistent for regression parameters, there are the following problems.

1. We can not obtain an explicit form of the CPN estimator depending on distributions of explanatory variables and errors.
2. Tedious calculations are required to obtain the explicit form of the CPN estimator.

Purpose of this research

Table: Bias correction based on the corrected score function

Explanatory variable	Normal	General
Univariate	Kukush et al. [2]	Not finished
Multivariate(All error)	Shklyar et al. [4]	Not finished
Multivariate(Partial error)	Not finished	Not finished

We propose the following extension to the Corrected Score (CS) estimator discussed in Kukush et al. [2] and Shklyar et al. [4].

1. Explanatory variables are partially subject to errors in the multivariate EIV.
2. Explanatory variable and error vectors follow arbitrary distributions.

Contributions of this research

Under multivariate EIV included errors partially in the explanatory variables,

it becomes possible to perform a consistent estimation for the regression parameters

even when the CPN estimator cannot be computed in an explicit form.

A Poisson regression model with multivariate regressor

We assume a Poisson regression model between the objective variable Y and the vector of explanatory variables $\boldsymbol{X} = (X_1, \ldots, X_{p+q})'$:

$$Y|\boldsymbol{X} \sim \mathrm{Po}\left(\exp\left(\boldsymbol{\beta}'\begin{pmatrix} 1 \\ \boldsymbol{X} \end{pmatrix}\right)\right), \ \boldsymbol{\beta} = (\beta_0, \ldots, \beta_{p+q})' \in \mathbb{R}^{p+q+1}. \tag{1}$$

Subvectors of explanatory variables

We define subvectors of $\boldsymbol{X}$ as:

$$\boldsymbol{X}_o = (X_1, \dots, X_p)', \quad \boldsymbol{X}_e = (X_{p+1}, \dots, X_{p+q})'. \tag{2}$$

While $\boldsymbol{X}_o$ represents a vector of explanatory variables that can be observed directly, $\boldsymbol{X}_e$ represents a vector of explanatory variables that cannot be observed directly.

Multivariate EIV

We assume here that

- $\boldsymbol{W} = \boldsymbol{X}_e + \boldsymbol{U}$.
- $\boldsymbol{U}$ is supposed to be independent of $(\boldsymbol{X}, Y)$.
- $(\boldsymbol{X}_i = (\boldsymbol{X}'_{o,i}, \boldsymbol{X}'_{e,i})', Y_i)$ $(i = 1, \ldots, n)$ are i.i.d. samples from the distribution of $(\boldsymbol{X}, Y)$ and $\boldsymbol{U}_i$ $(i = 1, \ldots, n)$ are i.i.d. samples from the distribution of $\boldsymbol{U}$.

Although we can observe $Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i = \boldsymbol{X}_{e,i} + \boldsymbol{U}_i$ $(i = 1, \ldots, n)$, we assume that $\boldsymbol{X}_e$ and $\boldsymbol{U}$ cannot be observed directly.

Assumptions of EIV model

In EIV models, it is common that some of the parameters of the model are known [6].

- Regression with known mean and variance of $\boldsymbol{U}$.
- Regression with known mean of $\boldsymbol{U}$ and variance ratio of components of $\boldsymbol{X}_e$ and $\boldsymbol{W}$.

In this study, we assume that the mean and variance of $\boldsymbol{U}$ are known.

Partial-error Naive (PN) estimator

PN estimator

The estimator using the observed variables $\boldsymbol{W}$ with errors in place of the unobserved explanatory variables $\boldsymbol{X}_e$ in the likelihood equation of $Y|\boldsymbol{X}$ [5].

Misspecified estimation of the PN estimator

In other words, the PN estimator applies MLE using a misspecified model that assumes

$$Y \left| \begin{pmatrix} \boldsymbol{X}_o \\ \boldsymbol{W} \end{pmatrix} \right. \sim \mathrm{Po}\left(\exp\left(\beta' \begin{pmatrix} 1 \\ \boldsymbol{X}_o \\ \boldsymbol{W} \end{pmatrix}\right)\right).$$

Inconsistency of the PN estimator

The PN estimator is inconsistent for $\boldsymbol{\beta}$ because it estimates the parameter incorrectly.

$$\hat{\boldsymbol{\beta}}^{(PN)} \quad \xrightarrow{a.s.} \quad \boldsymbol{b} \quad \neq \quad \boldsymbol{\beta},$$

where $\boldsymbol{b}$ is a constant vector defined as a solution to the following estimating equation:

$$\mathsf{E}_{\boldsymbol{X},\boldsymbol{W}}[\mathsf{E}_{Y|(\boldsymbol{X},\boldsymbol{W})}[\{Y-\exp(b_0+\boldsymbol{b}_1'\boldsymbol{X}_o+\boldsymbol{b}_2'\boldsymbol{W})\}(1,\boldsymbol{X}_o',\boldsymbol{W}')']] = \boldsymbol{0}_{p+q+1}.$$

Corrected Partial-error Naive (CPN) estimator

The PN estimator has an asymptotic bias for true β.

$\Rightarrow$ Wada and Kurosawa [5] proposed the CPN estimator as a consistent estimator by correcting the bias of the PN estimator.

Score function of a Poisson regression model

The score function of a Poisson regression model (1) is given by

$$\psi(\boldsymbol{\beta}|Y, \boldsymbol{X}) = (Y - \exp(\beta_0 + \boldsymbol{\beta}_1' \boldsymbol{X}_o + \boldsymbol{\beta}_2' \boldsymbol{X}_e)) \begin{pmatrix} 1 \\ \boldsymbol{X} \end{pmatrix}.$$

Corrected score function

A function such that its conditional expectation given Y and $\boldsymbol{X}$ equals the true score function in EIV models [7, 8].

Therefore, we construct the corrected score function $\psi_c : \mathbb{R}^{p+q+1} \to \mathbb{R}^{p+q+1}$ to satisfy

$$\mathbf{E}\left[\psi_c(\beta|Y, \boldsymbol{X}_o, \boldsymbol{W})|Y, \boldsymbol{X}\right] = \psi(\beta|Y, \boldsymbol{X}).$$

Corrected score function under multivariate EIV

The function $\psi_c(\boldsymbol{\beta}|Y, \boldsymbol{X}_o, \boldsymbol{W})$ can be obtained by

$$Y\begin{pmatrix} 1 \\ \boldsymbol{X}_o \\ \boldsymbol{W} - \mathbf{E}[\boldsymbol{U}] \end{pmatrix} - \frac{\exp(\beta_0 + \boldsymbol{\beta}_1'\boldsymbol{X}_o + \boldsymbol{\beta}_2'\boldsymbol{W})}{M_{\boldsymbol{U}}(\boldsymbol{\beta}_2)}\begin{pmatrix} 1 \\ \boldsymbol{X}_o \\ \boldsymbol{W} - \frac{1}{M_{\boldsymbol{U}}(\boldsymbol{\beta}_2)}\frac{\partial M_{\boldsymbol{U}}(\boldsymbol{\beta}_2)}{\partial \boldsymbol{\beta}_2} \end{pmatrix}.$$

Partial-error Corrected Score (PCS) estimator

Thus, we define the PCS estimator as a solution to the corrected estimating equation:

$$\sum_{i=1}^{n} \psi_c(\hat{\beta}^{(PCS)} | Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i) = \boldsymbol{0}_{p+q+1} \tag{3}$$

and propose it as a consistent estimator for β. This definition is a natural extension of the univariate CS estimator discussed in Kukush et al. [2].

Numerical algorithm

We derive a solution to equation (3) using the Newton-Raphson method with the following update formula.

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} - \left(\sum_{i=1}^{n} \frac{\partial \psi_c(\hat{\boldsymbol{\beta}}^{(k)} | Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i)}{\partial \boldsymbol{\beta}'} \right)^{-1} \times \left(\sum_{i=1}^{n} \psi_c(\hat{\boldsymbol{\beta}}^{(k)} | Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i) \right) \qquad (k = 0, 1, 2, \ldots),$$

where $\boldsymbol{\beta}^{(0)} = \hat{\boldsymbol{\beta}}^{(PN)}$.

Simulation studies

- Multivariate normal distribution + normal error $(p = 2, q = 1)$

$$\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \Sigma), \; U \sim N(0, \sigma^2)$$

- Gumbel distribution + normal error $(p = 0, q = 2)$

$$X_1 \sim Gu(\mu_1, \eta_1), \; X_2 \sim Gu(\mu_2, \eta_2), \; U_1 \sim N(0, \sigma_1^2), \; U_2 \sim N(0, \sigma_2^2)$$

We perform simulations in above 2 ways with n as the sample size and MC as the number of simulations.

Simulation configuration 1 ($p = 2, q = 1$)

We assume the Poisson regression model with multivariate EIV ($p = 2, q = 1$).

$$Y|\boldsymbol{X} \sim \mathrm{Po}\left(\exp\left(\boldsymbol{\beta}'\begin{pmatrix} 1 \\ \boldsymbol{X} \end{pmatrix}\right)\right),\ \boldsymbol{\beta} \in \mathbb{R}^4,$$

$$\boldsymbol{X}_o = (X_1, X_2)',\ X_e = X_3,\ W = X_3 + U,$$

$$\boldsymbol{X} = (X_1, X_2, X_3)' \sim N_3(\boldsymbol{\mu}, \Sigma),\ U \sim N(0, \sigma^2),$$

$$\boldsymbol{\mu} \in \mathbb{R}^3,\ \Sigma \in \mathbb{R}^{3\times 3},\ 0 < \sigma^2 < \infty.$$

Parameter configuration (Multivariate normal+normal error)

We set the parameters as

$$\beta = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.2 \\ 0.3 \end{pmatrix}, \ \mu = \begin{pmatrix} 1 \\ 1.2 \\ 0.5 \end{pmatrix}, \ \Sigma = \begin{pmatrix} 1 & 0.2 & -0.5 \\ 0.2 & 1.1 & 0.3 \\ -0.5 & 0.3 & 1.2 \end{pmatrix},$$

$n = 5000$, $MC = 10000$, $\epsilon = 10^{-8}$, the maximum number of iterations as 100 and calculate the estimated bias with $\sigma^2 = 0.25, 0.5, 0.75, 1, 1.25$.

Results of estimated bias (β_0, β_1)

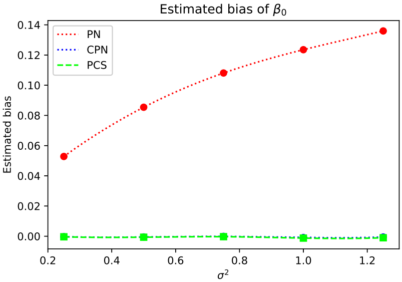

Figure: Estimated bias for β_0

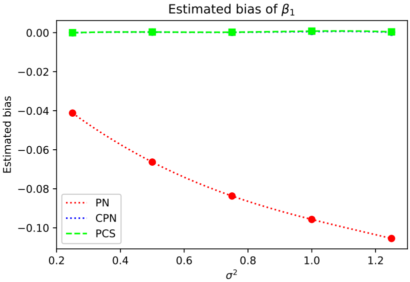

Figure: Estimated bias for β_1

Results of estimated bias (β_2, β_3)

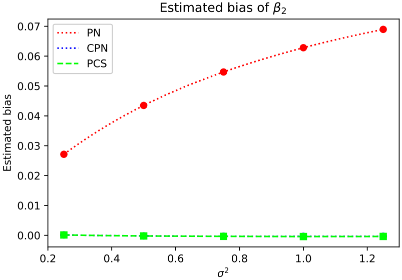

Figure: Estimated bias for β_2

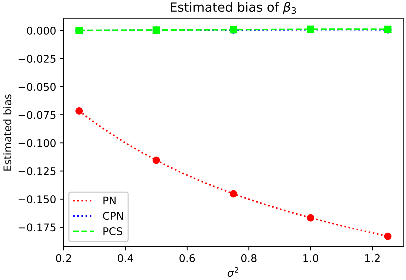

Figure: Estimated bias for β_3

Simulation configuration 2 ($p = 0, q = 2$)

We assume the Poisson regression model with multivariate EIV ($p = 0, q = 2$).

$$Y|\boldsymbol{X} \sim \mathrm{Po}\left(\exp\left(\boldsymbol{\beta}'\begin{pmatrix} 1 \\ \boldsymbol{X} \end{pmatrix}\right)\right), \ \boldsymbol{\beta} \in \mathbb{R}^3,$$

$$\boldsymbol{X}_e = (X_1, X_2)' = \boldsymbol{X}, \ W_1 = X_1 + U_1, \ W_2 = X_2 + U_2,$$

$$X_1 \sim Gu(\mu_1, \eta_1), \ X_2 \sim Gu(\mu_2, \eta_2), \ U_1 \sim N(0, \sigma_1^2), \ U_2 \sim N(0, \sigma_2^2),$$

$$\mu_1, \mu_2 \in \mathbb{R}, \ \eta_1 > 0, \ \eta_2 > 0, \ 0 < \sigma_1^2 < \infty, \ 0 < \sigma_2^2 < \infty.$$

Parameter configuration (Gumbel+normal error)

We set the parameters as

$$\beta = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.3 \end{pmatrix}, \ \mu_1 = -0.5, \ \eta_1 = 1, \ \mu_2 = 0.25, \ \eta_2 = 0.5, \ \sigma_2^2 = 0.5,$$

$n = 5000, \ MC = 10000, \ \epsilon = 10^{-8}$, the maximum number of iterations as 100 and calculate the estimated bias with $\sigma_1^2 = 0.25, 0.5, 0.75, 1, 1.25$.

Results of estimated bias (β_0, β_1)

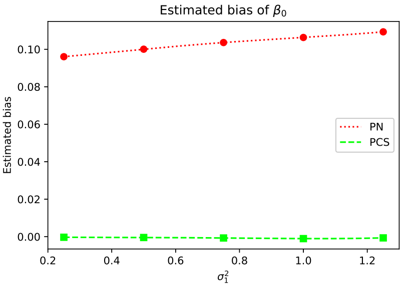

Figure: Estimated bias for β_0

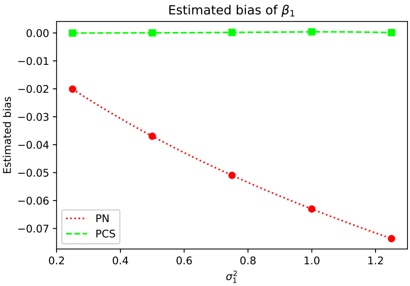

Figure: Estimated bias for β_1

Results of estimated bias (β_2)

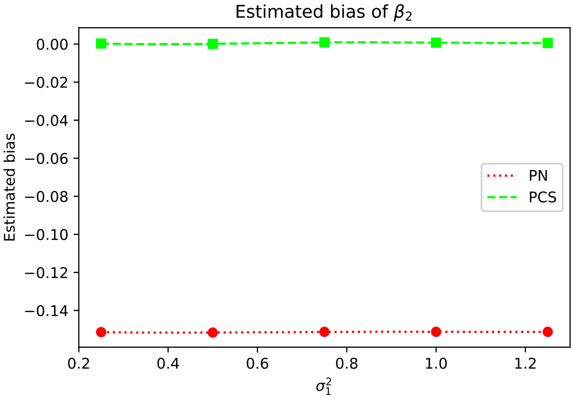

Figure: Estimated bias for β_2

Summary

- We extend the CS estimator to multivariate EIV with partially included errors in explanatory variables.
- Consistent estimation is possible even when the CPN estimator cannot be computed in closed form.
- The PCS estimator achieves estimation performance comparable to the CPN estimator, even in cases where the CPN estimator is computable.

Future works

- Consider numerical algorithms with reduced computational cost.
- Derive an asymptotic distribution of the PCS estimator.
- Construct confidence intervals of the regression parameter under multivariate EIV.

Acknowledgements

This research was supported by JSPS KAKENHI Grant Number JP25K07136.

References I

[1] G.T.Goldman, J.A.Mulholland, A.G.Russell, M.J.Strickland, M.Klein, L.A.Waller, and P.E.Tolbert.
Impact of exposure measurement error in air pollution epidemiology: effect of error type in time-series studies.
Environmental Health, 10(61), 2011.

[2] A.Kukush, H.Schneeweiss, and R.Wolf.
Three estimators for the Poisson regression model with measurement errors.
Statistical Papers, 45(3):351–368, 2004.

[3] K.Wada and T.Kurosawa.
The naive estimator of a Poisson regression model with a measurement error.
Journal of Risk and Financial Management, 16(3):16–30, 2023.

References II

[4] S.Shklyar and H.Schneeweiss.
A comparison of asymptotic covariance matrices of three consistent estimators in the Poisson regression model with measurement errors.
Journal of Multivariate Analysis, 94(2):250–270, 2005.

[5] K.Wada and T.Kurosawa.
Bias correction of partial-error in variables in a poisson regression model.
Statistical Theory and Related Fields, 2025.
to appear.

[6] W.A.Fuller.
Measurement Error Models.
Wiley, Hoboken, 1987.

References III

[7] T.Nakamura.
Corrected score function for errors-in-variables models:methodology and application to generalized linear models.
Biometrika, 77(1):127–137, 1990.

[8] L.A.Stefanski.
Unbiased estimation of a nonlinear function a normal mean with application to measurement error models.
Communications in Statistics - Theory and Methods, 18(12):4335–4358, 1989.

[9] A.Kukush and S.Shklyar.
Comparison of three estimators in poisson errors-in-variables model with one covariate.
Sonderforschungsbereich, 386:Paper 293, 2002.

Notations

$M_{\boldsymbol{X}}$	:	Moment generating function of a random vector $\boldsymbol{X}$
$K_{\boldsymbol{X}}$	:	Cumulant generating function of a random vector $\boldsymbol{X}$
$\boldsymbol{\beta}_1$	:	Subvector of $\boldsymbol{\beta}$ corresponding to $\boldsymbol{X}_o$
$\boldsymbol{\beta}_2$	:	Subvector of $\boldsymbol{\beta}$ corresponding to $\boldsymbol{X}_e$

Definition of the PN estimator

The PN estimator $\hat{\beta}^{(PN)}$ for β is defined as the solution of the following equation [5]:

$$S_n(\hat{\beta}^{(PN)}) = \mathbf{0}_{p+q+1},$$

where

$$\begin{aligned} S_n(\tilde{\boldsymbol{b}}) &= \frac{1}{n}\sum_{i=1}^{n}\{Y_i - \exp(\tilde{b}_0 + \tilde{\boldsymbol{b}}_1'\boldsymbol{X}_{o,i} + \tilde{\boldsymbol{b}}_2'\boldsymbol{W}_i)\}(1, \boldsymbol{X}_{o,i}', \boldsymbol{W}_i')', \\ \tilde{\boldsymbol{b}} &= (\tilde{b}_0, \ldots, \tilde{b}_{p+q})', \end{aligned}$$

$\tilde{\boldsymbol{b}}_1 = (\tilde{b}_1, \ldots, \tilde{b}_p)'$, $\tilde{\boldsymbol{b}}_2 = (\tilde{b}_{p+1}, \ldots, \tilde{b}_{p+q})'$ are subvectors of $\tilde{\boldsymbol{b}}$ and $\mathbf{0}_{p+q+1}$ is a $(p+q+1)$-dimensional vector with zeros.

Exact distribution of Y given (X_o, W)

$$
\begin{aligned}
& f_{Y|(\boldsymbol{X}_o,\boldsymbol{W})}(y|(\boldsymbol{x}_o,\boldsymbol{w})) \\
= & \frac{\int \mathrm{Po}\left(\exp\left(\beta'\begin{pmatrix}1\\ \boldsymbol{x}\end{pmatrix}\right)\right) f_U(\boldsymbol{w}-\boldsymbol{x}_e) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}_e}{\int f_U(\boldsymbol{w}-\boldsymbol{x}_e) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}_e}.
\end{aligned}
$$

The exact distribution of Y given $(\boldsymbol{X}_o, \boldsymbol{W})$ is not a Poisson regression model.

Assumptions for the CPN estimator

We assume the following conditions (C1) and (C2).

- (C1) $M_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{pmatrix}$, $M_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{pmatrix}$, and $M_{\boldsymbol{U}}(\boldsymbol{b}_2)$ exist.
- (C2) $\det \dfrac{\partial \boldsymbol{G}}{\partial\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{pmatrix}'} \neq 0$ is satisfied where

$$\boldsymbol{G} = \begin{pmatrix} \frac{\partial}{\partial \boldsymbol{b}_1} K_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{pmatrix} - \frac{\partial}{\partial \boldsymbol{\beta}_1} K_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{pmatrix} \\ \frac{\partial}{\partial \boldsymbol{b}_2} K_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{pmatrix} + \frac{\partial}{\partial \boldsymbol{b}_2} K_{\boldsymbol{U}}(\boldsymbol{b}_2) - \mathbf{E}[\boldsymbol{U}] - \frac{\partial}{\partial \boldsymbol{\beta}_2} K_{\boldsymbol{X}}\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{pmatrix} \end{pmatrix}.$$

Formula of the CPN estimator

$$\hat{\beta}_0^{(CPN)} = \hat{\beta}_0^{(PN)} + \log\left(\frac{M_{\boldsymbol{X}} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1^{(PN)} \\ \hat{\boldsymbol{\beta}}_2^{(PN)} \end{pmatrix} M_{\boldsymbol{U}}(\hat{\boldsymbol{\beta}}_2^{(PN)})}{M_{\boldsymbol{X}} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1^{(CPN)} \\ \hat{\boldsymbol{\beta}}_2^{(CPN)} \end{pmatrix}} \right),$$

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_1^{(CPN)} \\ \hat{\boldsymbol{\beta}}_2^{(CPN)} \end{pmatrix} = \boldsymbol{h} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1^{(PN)} \\ \hat{\boldsymbol{\beta}}_2^{(PN)} \end{pmatrix},$$

where $\boldsymbol{h}$ is a continuously differentiable implicit function with $\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = \boldsymbol{h} \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix}$ in the neighborhood of $\left(\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} \right)$ satisfying $\boldsymbol{G} = \boldsymbol{0}$.

Application example of the CPN estimator

We assume the Poisson regression model with multivariate EIV $(p = 2, q = 1)$.

$$Y|\boldsymbol{X} \sim \mathrm{Po}\left(\exp\left(\boldsymbol{\beta}'\begin{pmatrix} 1 \\ \boldsymbol{X} \end{pmatrix}\right)\right),\ \boldsymbol{\beta} \in \mathbb{R}^4,$$

$$\boldsymbol{X}_o = (X_1, X_2)',\ X_e = X_3,\ W = X_3 + U,$$

$$\boldsymbol{X} = (X_1, X_2, X_3)' \sim N_3(\boldsymbol{\mu}, \Sigma),\ U \sim N(0, \sigma^2),$$

$$\boldsymbol{\mu} \in \mathbb{R}^3,\ \Sigma \in \mathbb{R}^{3\times 3},\ 0 < \sigma^2 < \infty.$$

Application example of the CPN estimator

We use following partition expressions of $\boldsymbol{\mu}$ and Σ.

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \mu_3 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_1 & \boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_3' & \sigma_{33} \end{pmatrix},$$

where

$$\boldsymbol{\mu}_1 = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \Sigma_1 = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}, \ \boldsymbol{\sigma}_3 = \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix}.$$

$$
\begin{aligned}
\hat{\beta}_0^{(CPN)} &= \hat{\beta}_0^{(PN)} + \left(\hat{\boldsymbol{\beta}}_1^{(PN)} - \hat{\boldsymbol{\beta}}_1^{(CPN)}\right)' \boldsymbol{\mu}_1 + \left(\hat{\beta}_3^{(PN)} - \hat{\beta}_3^{(CPN)}\right) \mu_3 \\
&\quad - \frac{1}{2}\left(\hat{\boldsymbol{\beta}}_1^{(CPN)'} \Sigma_1 \hat{\boldsymbol{\beta}}_1^{(CPN)} + \hat{\beta}_3^{(CPN)}\left(2\boldsymbol{\sigma}_3' \hat{\boldsymbol{\beta}}_1^{(CPN)} + \sigma_{33}\hat{\beta}_3^{(CPN)}\right)\right) \\
&\quad + \frac{1}{2}\left(\hat{\boldsymbol{\beta}}_1^{(PN)'} \Sigma_1 \hat{\boldsymbol{\beta}}_1^{(PN)} + \hat{\beta}_3^{(PN)}\left(2\boldsymbol{\sigma}_3' \hat{\boldsymbol{\beta}}_1^{(PN)} + \sigma_{33}\hat{\beta}_3^{(PN)}\right)\right) \\
&\quad + \frac{1}{2}\sigma^2 \hat{\beta}_3^{(PN)2}, \\
\hat{\boldsymbol{\beta}}_1^{(CPN)} &= \hat{\boldsymbol{\beta}}_1^{(PN)} - \frac{\sigma^2 \hat{\beta}_3^{(PN)}}{\sigma_{33} - \boldsymbol{\sigma}_3' \Sigma_1^{-1} \boldsymbol{\sigma}_3} \Sigma_1^{-1} \boldsymbol{\sigma}_3, \\
\hat{\beta}_3^{(CPN)} &= \frac{\sigma_{33} + \sigma^2 - \boldsymbol{\sigma}_3' \Sigma_1^{-1} \boldsymbol{\sigma}_3}{\sigma_{33} - \boldsymbol{\sigma}_3' \Sigma_1^{-1} \boldsymbol{\sigma}_3} \hat{\beta}_3^{(PN)}.
\end{aligned}
$$

Estimation of μ and Σ

We can estimate $\boldsymbol{\mu}$ and Σ in the formula of the CPN estimator using the method of moments in terms of $\boldsymbol{X}_o = (X_1, X_2)'$ and W.

$$\hat{\boldsymbol{\mu}}_1 = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_{o,i}, \ \hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n} w_i,$$

$$\hat{\Sigma}_1 = \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{x}_{o,i} - \bar{\boldsymbol{x}}_o)(\boldsymbol{x}_{o,i} - \bar{\boldsymbol{x}}_o)',$$

$$\hat{\boldsymbol{\sigma}}_3 = \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{x}_{o,i} - \bar{\boldsymbol{x}}_o)(w_i - \bar{w}),$$

$$\hat{\sigma}_{33} = \frac{1}{n}\sum_{i=1}^{n} (w_i - \bar{w})^2 - \sigma^2.$$

PCS estimator

Theorem 1

Let $Y|\boldsymbol{X}$ be a Poisson regression model in (1) with (2). We assume the following condition (D1).

- **D1** $M_{\boldsymbol{U}}(\boldsymbol{\beta}_2)$ *exists and is differentiable in the neighborhood of* $\boldsymbol{\beta}_2$.

Then, the PCS estimator $\hat{\boldsymbol{\beta}}^{(PCS)}$ of $\boldsymbol{\beta}$, which is defined as a solution of equation (3), is strongly consistent.

Convergence of the corrected estimating equation

By the strong law of large numbers, we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\psi_c(\hat{\beta}^{(PCS)}|Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i) \xrightarrow{a.s.} \mathsf{E}\left[\psi_c(\hat{\beta}^{(PCS)}|Y, \boldsymbol{X}_o, \boldsymbol{W})\right].$$

Expectation of the corrected score function

We can calculate the expectation of ψ_c as

$$\begin{aligned}
&\mathsf{E}_{\boldsymbol{X},Y,\boldsymbol{U}}[\psi_c(\boldsymbol{\beta}|Y,\boldsymbol{X}_o,\boldsymbol{W})] \\
=&\ \mathsf{E}_{\boldsymbol{X},Y}[\mathsf{E}[\psi_c(\boldsymbol{\beta}|Y,\boldsymbol{X}_o,\boldsymbol{W})|\boldsymbol{X},Y]] \\
=&\ \mathsf{E}_{\boldsymbol{X},Y}\left[(Y-\exp(\beta_0+\boldsymbol{\beta}_1'\boldsymbol{X}_o+\boldsymbol{\beta}_2'\boldsymbol{X}_e))\begin{pmatrix}1\\ \boldsymbol{X}\end{pmatrix}\right]=\mathbf{0}.
\end{aligned}$$

Strong consistency of the PCS estimator

From the same argument in Kukush and Shklyar [9], we obtain the strong consistency of the PCS estimator:

$$\hat{\beta}^{(PCS)} \quad \xrightarrow{a.s.} \quad \beta.$$

Conditions for the convergence of numerical algorithm

We implicitly assume the following conditions (D2)-(D4) to ensure that the algorithms converge.

- (D2) $\mathbf{E}\left[\frac{\partial \psi_c(\boldsymbol{\beta}|Y,\boldsymbol{X}_o,\boldsymbol{W})}{\partial \boldsymbol{\beta}'}\right]$ exists and is nonsingular.
- (D3) ψ_c is differentiable and locally Lipschitz continuous.
- (D4) Satisfying

$$\alpha_0 = M \left\| \left(\frac{\partial F(\boldsymbol{\beta}^{(0)})}{\partial \boldsymbol{\beta}'} \right)^{-1} \right\| \left\| \boldsymbol{h}^{(0)} \right\| \le \frac{1}{2},$$

  where

$$\boldsymbol{h}^{(0)} = -\left(\frac{\partial F(\boldsymbol{\beta}^{(0)})}{\partial \boldsymbol{\beta}'} \right)^{-1} F(\boldsymbol{\beta}^{(0)}), \quad F(\boldsymbol{\beta}) = \sum_{i=1}^{n} \psi_c(\boldsymbol{\beta}|Y_i, \boldsymbol{X}_{o,i}, \boldsymbol{W}_i)$$

  and M is a Lipschitz constant of $\frac{\partial F}{\partial \boldsymbol{\beta}'}$.

Stopping criterion for numerical algorithm

The iteration is stopped when one of the following stopping criteria is fulfilled.

1. $||\boldsymbol{\beta}^{(k+1)} - \boldsymbol{\beta}^{(k)}|| < \epsilon$ is satisfied for a pre-specified $\epsilon > 0$.
2. The maximum number of iterations has been reached.