Metrika
https://doi.org/10.1007/s00184-018-0664-1

Assessing the robustness of estimators when fitting Poisson inverse Gaussian models

Kimberly S. Weems[1] · **Paul J. Smith**[2]

Received: 24 October 2017

Abstract The generalized linear mixed model (GLMM) extends classical regression analysis to non-normal, correlated response data. Because inference for GLMMs can be computationally difficult, simplifying distributional assumptions are often made. We focus on the robustness of estimators when a main component of the model, the random effects distribution, is misspecified. Results for the maximum likelihood estimators of the Poisson inverse Gaussian model are presented.

Keywords Poisson mixed models · Inverse Gaussian distribution · Influence function · Directional derivative · Maximum likelihood estimation

1 Introduction

Poisson mixed models, a class of generalized linear mixed models (McCulloch et al. 2008), are often used to analyze count data that exhibit overdispersion. For example, see Dean and Nielsen (2007), Karlis and Xekalaki (2005), Ven and Weber (1995), and Hougaard et al. (1997). For these models, we assume that the conditional distribution of the response follows a Poisson distribution with a random mean. The mean incorporates the random effects used to model the overdispersion.

This work was supported by National Science Foundation Grant #1700235. Additional support was provided by a grant from North Carolina Central university.

✉ Kimberly S. Weems
ksweems@nccu.edu

Paul J. Smith
pjs@math.und.edu

1 Department of Mathematics and Physics, North Carolina Central University, Durham, NC, USA

2 Department of Mathematics, University of Maryland, College Park, MD, USA

Published online: 04 June 2018

When estimating the model parameters via maximum likelihood, the distribution of the random effects, known as the mixing distribution, is often chosen for computational convenience. This assumption motivated us to consider the robustness of the maximum likelihood estimators (MLEs) of the fixed effects and of the random effects variance when the mixing distribution is slightly contaminated. In particular, the mixing distribution is presumed to be an ϵ-contamination of a specified parametric family. This contamination results in misspecification of the marginal distribution of the response variable.

Some research has indicated that small perturbations of the random effects distribution have minimal impact on parameter estimation. For instance, see McCulloch and Neuhaus (2013), Heagerty and Zeger (2000) and Neuhaus et al. (1992). Yet, other research points to greater sensitivity of estimators, including Heckman and Singer (1984) and Litière et al. (2008), among others.

We examine the effects of this misspecification using an infinitesimal approach based on the influence function (Hampel et al. 1986 and Huber 1981). Gustafson (1996) followed the influence function approach to consider the robustness of MLEs for certain conjugate mixture models under mixing distribution misspecification, and Weems and Smith (2004) extended this approach to include a regression structure in the mean. Specifically, Weems and Smith (2004) considered an influence function for mixed Poisson regression models that can be used to assess the effects of mixing distribution misspecification on MLEs for regression parameters and the variance component. Using saddlepoint techniques of Bruijn (1953), they showed that the integral of the influence function for MLEs of the Poisson-lognormal regression model is uniformly bounded over a certain class of distribution functions; hence, the MLEs are robust against mixing distribution misspecification. More recently, Verbeke and Molenberghs (2013) considered the gradient function (directional derivative) as a qualitative goodness-of-fit assessment of the random effects distribution.

Our focus is the Poisson inverse Gaussian model. In Sect. 2, we review the influence function approach of Gustafson (1996) and Weems and Smith (2004) for determining the robustness of MLEs in mixed Poisson regression models. Section 3 focuses on the robustness of MLEs for the Poisson inverse Gaussian model. A simulation study is presented in Sect. 4, and a summary is given in Sect. 5.

2 Robustness concepts

2.1 Influence function

Let Ψ be an estimating function, and define the functional $T(F)$ to be the solution, in θ, of

$$\int \Psi(x;\theta)F(dx) = 0 \tag{1}$$

for a p-dimensional parameter vector θ. As discussed by Gustafson (1996) and Weems and Smith (2004), $T = T(X_1, X_2, \ldots, X_n)$ can be regarded as a function $T(F_n)$ applied to the empirical cdf of the data $X = (X_1, X_2, \ldots, X_n)$. In this case, $T(F_n)$ estimates $T(F)$, where F is the true distribution of the data.

If the distribution of X is perturbed from F to $F_\epsilon = (1-\epsilon)F + \epsilon G$, then one may consider

$$\dot{T}(F;G) = (\partial/\partial\epsilon)T(F_\epsilon)|_{\epsilon=0} \tag{2}$$

as a measure of the sensitivity of T to small perturbations of F. The quantity (2) is the Gâteaux derivative of $T(F)$ in the direction of G and is found by implicit differentiation of the equation

$$\int \Psi(x; T(F_\epsilon))[(1-\epsilon)F + \epsilon G](dx) = 0,$$

which gives

$$\dot{T}(F;G) = \left[-\int \nabla_\theta \Psi(x; T(F))F(dx)\right]^{-1} \left[\int \Psi(x; T(F))G(dx)\right].$$

Using this quantity, we state the following definition of robustness.

Definition 1 Let T be an M-estimate. Suppose that the distribution function F is contaminated by an epsilon amount of a distribution G. Then T is robust against distribution misspecification if

$$\left[-\int \nabla_\theta \Psi(x; T(F))F(dx)\right]^{-1} \left[\int \Psi(x; T(F))G(dx)\right] \tag{3}$$

is bounded for all G.

It is appropriate to examine the quantity given in (3) because it gives a first-order approximation to the asymptotic bias in estimating θ that is introduced by the ϵ-contamination of F by a distribution G.

2.2 Misspecificiation in Poisson mixed models

Let Y_i, U_i, and X_i denote the response variables, random effects and covariates, respectively, for $i = 1, \ldots, n$. Suppose the conditional distribution of $Y_i|X_i$ is Poisson with mean $U_i\mu_i$, where $\mu_i = \mu(X_i) = \exp(\beta_0 + \beta_1 X_i)$, and U_i and X_i are independent. Note that β_0 and β_1 are unknown regression parameters.

Let F denote the nominal distribution function of the random effects U_i and f denote the corresponding density function. The mean and variance of U_i are 1 and τ, respectively. Using maximum likelihood, our goal is to estimate $\theta = (\beta_0, \beta_1, \tau)$. The marginal density of Y_i is

$$P(Y_i = y_i; X_i, \theta) = \int_0^\infty \frac{(u\mu_i)^{y_i} \exp(-u\mu_i)}{y_i!} f(u)du, \tag{4}$$

with marginal mean and variance of μ_i and $\mu_i(1+\mu_i\tau)$, respectively. This model includes a single covariate X_i; however, the extension to multiple covariates is straightforward.

To determine the MLE $\hat{\theta}$, we maximize $l(\theta; \mathbf{Y}) = \log(L(\theta; \mathbf{Y}))$, where $\mathbf{Y} = (Y_1, \ldots, Y_n)$ and

$$\log(L(\theta; \mathbf{Y})) = \prod_{i=1}^{n} P(Y_i = y_i; X_i, \theta).$$

Let $\Psi = \nabla_\theta l(\theta; \mathbf{Y})$. Then (3) becomes $\int_0^\infty \{I^{-1}(\theta) s(\theta; u)\} G(du)$, where

$$I(\theta) = -\mathbb{E}\left[\nabla_\theta^{\otimes 2} l(\theta; \mathbf{Y})\right]$$

is the Fisher information matrix,

$$s(\theta; u) = \mathbb{E}\left[\nabla_\theta l(\theta; \mathbf{Y}) | \, U_i = u\right]$$

is the conditional expected score matrix, and all expectations are taken with respect to the nominal distribution F. The integrand $I^{-1}(\theta)s(\theta; u)$ is an influence function for misspecification of the mixing distribution F (Gustafson 1996). Hence, integrating this quantity captures the effect of a contaminating distribution G on the MLE of θ. We let

$$\mathbf{IF}(u; \hat{\theta}, F) = \left[IF(u; \hat{\beta}_0, F), IF(u; \hat{\beta}_1, F), IF(u; \hat{\tau}, F)\right]^T$$

represent the 3×1 matrix of influence functions for the individual parameter estimates. Hence, restating Definition 1, if

$$\int \mathbf{IF}(x; T, F) G(dx) = \int_0^\infty \{I^{-1}(\theta) s(\theta; u)\} G(du) \tag{5}$$

is bounded for all G, then $\hat{\theta}$ is robust. We note that if $I^{-1}(\theta)$ is well-behaved then bounding (5) reduces to bounding the integral of the conditional expected score matrix. We make the following assumptions:

1. Let $\mathcal{F} = \{F | F \text{ is a cdf on } (0, \infty), \int uF(du) = 1, \text{ and } \int u^2 F(du) = 1 + \tau\}$. We let the nominal distribution $F \in \mathcal{F}$ and, likewise, the contaminating distribution $G \in \mathcal{F}$.
2. $n^{-1} I \longrightarrow I^*$ as $n \longrightarrow \infty$, where I^* is positive definite.
3. The covariates $X_1, X_2, \ldots, X_n$ are an i.i.d. sample from a nondegenerate distribution whose support is a compact region of $\mathbb{R}^p$. Additionally, $\mathrm{Cov}(X)$ is positive definite, where $X = (X_1, X_2, \ldots, X_n)^T$.
4. The interchange of expectation and differentiation of the log-likelihood and its derivatives is permitted.

These assumptions are similar to those presented in Gustafson (1996). In particular, Assumptions 1, 2, and 3 ensure identifiability of the parameters, and Assumption 4 follows usual regularity conditions for asymptotic results (see, for instance, Bickel and Doksum 2001).

There are many approaches to robustness beyond influence function methods, such as the breakdown point. In a regression context, which is the focus of this paper, there are even more ways to assess robustness, such as those found in Bickel (1984) and Rieder (1994). However, our scope is limited to Definition 1 above focused on infinitesimal misspecification of the distribution of the random effects.

In the next section, we use an influence function approach to determine the effect of G on the MLEs for the fixed effect parameters β_0 and β_1 and the random effects parameter τ when fitting a Poisson inverse Gaussian regression model. We suppress subscripts to simplify notation.

3 Poisson inverse Gaussian model

Consider the Poisson inverse Gaussian model, where we denote the nominal mixing distribution F as IG$(1, 1/\tau)$. The corresponding density is given by

$$f(u) = \frac{1}{(2\pi\tau u^3)^{1/2}} \exp\left\{-\frac{(u-1)^2}{2\tau u}\right\}, \tag{6}$$

$u > 0$, $\tau > 0$. The Poisson inverse Gaussian distribution has been studied extensively by Ong (1998), Shaban (1981), Seshadri (1993) and Dean et al. (1989). In addition, Seshadri (1999) points out several applications in actuarial science, linguistics, and ecology. Zha et al. (2016) and Shoukri et al. (2004) use the Poisson inverse Gaussian regression model to analyze motor vehicle crashes and disease incidence, respectively.

The marginal probabilities of a Poisson inverse Gaussian mixture, denoted by PIG, are as follows:

$$\Pr(Y = y|X, \theta) = \frac{2\mu^y \exp(1/\tau)}{y!(2\pi\tau)^{1/2}} K_{y-\frac{1}{2}}\left[\sqrt{\tau^{-1}(\tau^{-1}+2\mu)}\right]\left(\frac{1}{1+2\mu\tau}\right)^{\frac{1}{2}(y-\frac{1}{2})}, \tag{7}$$

where

$$K_k(\nu) = \frac{1}{2}\int_0^\infty u^{k-1} \exp\left\{-\frac{\nu}{2}\left(u+\frac{1}{u}\right)\right\} du$$

denotes the modified spherical Bessel function of the third kind of order k (Abramowitz and Stegun 1972).

3.1 Asymptotic behavior of probability ratios

We present a theorem by Willmot (1990) that is used to gain some insight into the marginal Poisson inverse Gaussian probabilities. This theorem relates the behavior of the right tail probabilities to the corresponding right tail of the random effects distribution. We first introduce the following definition.

Definition 2 A positive function L, defined on $[0, \infty)$, varies slowly at infinity if, for all $c > 0$, $\lim_{x\to\infty} L(cx)/L(x) = 1$.

Let P_y denote the marginal probabilities of any mixed Poisson distribution, $y = 0, 1, 2, \ldots$ The notation $a(x) \sim b(x)$, $x \to \infty$, means $\lim_{x\to\infty} a(x)/b(x) = 1$.

Theorem 1 (Willmot 1990) *Let P_y denote probabilities of a mixed Poisson distribution so that*

$$P_y = \int_0^\infty \frac{(\lambda x)^y \exp(-\lambda x)}{y!} f(x)dx.$$

Suppose that

$$f(x) \sim C(x)x^\alpha \exp(-\beta x), \quad x \to \infty,$$

where $C(x)$ is a locally bounded function on $(0, \infty)$ which varies slowly at infinity, $\beta \geq 0$, and $-\infty < \alpha < \infty$ (with $\alpha < -1$ if $\beta = 0$). Then P_y satisfies

$$P_y \sim \frac{C(y)}{(\lambda + \beta)^{\alpha+1}} \left(\frac{\lambda}{\lambda + \beta}\right)^y y^\alpha, \quad y \to \infty.$$

We now apply this theorem to the PIG$(1, 1/\tau)$ distribution.

Proposition 1 *For the PIG$(1, 1/\tau)$ density with probabilities P_y*

$$P_y \sim \exp\left\{\frac{2y - 1}{2\tau y}\right\} \sqrt{\frac{\mu + (2\tau)^{-1}}{2\pi\tau}} \left(\frac{\mu}{\mu_i + (2\tau)^{-1}}\right)^y y^{-3/2}.$$

Proof Through algebraic manipulations, the IG$(1, 1/\tau)$ density in (6) can be rewritten as

$$f(u) = C(u)u^{-3/2} \exp\left(-\frac{u}{2\tau}\right),$$

where $C(u) = (\sqrt{2\pi\tau})^{-1/2} \exp\{(2u - 1)/2\tau u\}$. It is clear that $C(u)$ is locally bounded on $(0, \infty)$. Notice that for all $c > 0$,

$$\lim_{u\to\infty} \frac{C(cu)}{C(u)} = \lim_{u\to\infty} \exp\left(\frac{c - 1}{2\tau c u}\right) = 1,$$

so by definition $C(u)$ varies slowly at infinity. Therefore, by Theorem 1,

$$\begin{aligned} P_y &\sim \sqrt{2\pi\tau} \exp\left\{\frac{2y - 1}{2\tau y}\right\} \frac{1}{(\mu + (2\tau)^{-1})^{-3/2+1}} \left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)^y y^{-3/2} \\ &= \exp\left\{\frac{2y - 1}{2\tau y}\right\} \sqrt{\frac{\mu + (2\tau)^{-1}}{2\pi\tau}} \left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)^y y^{-3/2}. \end{aligned} \tag{8}$$

□

We note that this result is closely related to one by Teugels and Willmot (1987) which examines the asymptotic behavior of the probabilities for a different parameterization of the model.

In the following section, we will be concerned with the ratios P_{y+1}/P_y. The next proposition gives an asymptotic result for these ratios.

Proposition 2 *For the PIG*$(1, 1/\tau)$ *distribution*

$$\lim_{y\to\infty} \frac{P_{y+1}}{P_y} = \frac{2\mu\tau}{2\mu\tau + 1}.$$

Proof Let

$$l(y) = \exp\left\{\frac{2y-1}{2\tau y}\right\}\sqrt{\frac{\mu + (2\tau)^{-1}}{2\pi\tau}}\left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)^{y} y^{-3/2}.$$

From Proposition 1, $P_y \sim l(y)$. We have that

$$\begin{aligned}\lim_{y\to\infty} \frac{P_{y+1}}{P_y} &= \lim_{y\to\infty} \frac{P_{y+1}}{l(y+1)}\frac{l(y+1)}{P_y}\\ &= \lim_{y\to\infty} \frac{P_{y+1}}{l(y+1)}\frac{l(y+1)}{P_y}\frac{y^{-3/2}}{y^{-3/2}}\frac{\exp\{-1/(2\tau y)\}}{\exp\{-1/(2\tau y)\}}.\end{aligned} \tag{9}$$

Notice that

$$\begin{aligned}&l(y+1)\frac{y^{-3/2}}{y^{-3/2}}\frac{\exp\{-1/(2\tau y)\}}{\exp\{-1/(2\tau y)\}}\\ &= \exp\left\{\frac{(2y+1)}{2\tau(y+1)}\right\}\sqrt{\frac{\mu + (2\tau)^{-1}}{2\pi\tau}}\left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)^{y+1} (y+1)^{-3/2}\\ &\quad\times \frac{y^{-3/2}}{y^{-3/2}}\frac{\exp\{-1/(2\tau y)\}}{\exp\{-1/(2\tau y)\}}\\ &= \exp\left\{\frac{1}{\tau} - \frac{1}{2\tau y}\right\}\sqrt{\frac{\mu + (2\tau)^{-1}}{2\pi\tau}}\left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)^{y} y^{-3/2}\\ &\quad\times \exp\left\{\frac{1}{2\tau y} - \frac{1}{2\tau(y+1)}\right\}\left(\frac{\mu}{\mu + (2\tau)^{-1}}\right)\left(1 + \frac{1}{y}\right)^{-3/2}\\ &= l(y)\exp\left(\frac{1}{4\tau^2 y(y+1)}\right)\left(\frac{2\mu\tau}{2\mu\tau + 1}\right)\left(1 + \frac{1}{y}\right)^{-3/2}.\end{aligned}$$

Therefore, (9) becomes

$$\lim_{y\to\infty} \frac{P_{y+1}}{l(y+1)} \times \lim_{y\to\infty} \frac{l(y)}{P_y} \times \lim_{y\to\infty} \exp\left(\frac{1}{4\tau^2 y(y+1)}\right)\left(\frac{2\mu\tau}{2\mu\tau+1}\right)\left(1+\frac{1}{y}\right)^{-3/2}$$
$$= \frac{2\mu\tau}{2\mu\tau+1}.$$

□

3.2 Robustness of maximum likelihood estimators

We examine the robustness of the MLEs for β_0, β_1, and τ. Recall that $\mu = \mu(X)$. Let

$$V = V(X) = \mu - \mu^2(1+\tau) + \mathbb{E}_{Y|X}\left[(Y+1)\frac{P_{Y+1}}{P_Y}\right]^2$$

and

$$\kappa = \kappa(X) = \frac{(1+\tau\mu)}{\tau^2}\left(\frac{\mu}{1+\tau\mu} - V\right).$$

Then, the influence function for the PIG$(1, 1/\tau)$ is given by

$$\mathbf{IF}(u; \hat{\theta}, F) = I^{-1}(\theta)s(\theta; u), \tag{10}$$

where

$$I(\theta) = \mathbb{E}_X \begin{pmatrix} V & \cdot & \cdot \\ XV & X^2V & \cdot \\ \kappa & X\kappa & -\kappa/\mu^2 \end{pmatrix} \tag{11}$$

and

$$s(\theta; u) = \mathbb{E}_X \begin{pmatrix} \mathbb{E}_{Y|X,U}\left[Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right] \\ X\,\mathbb{E}_{Y|X,U}\left[Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right] \\ \left(\frac{1+\tau\mu}{\tau^2}\right)\mathbb{E}_{Y|X,U}\left[\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y} - 1\middle| X, U = u\right] \end{pmatrix}. \tag{12}$$

Details concerning the computation of the Fisher information matrix $I(\theta)$ can be found in Dean et al. (1989).

Recall that the integral of the influence function should be bounded in order for the estimators to be robust. Notice that the influence function for PIG$(1, 1/\tau)$ is a linear combination of the terms of the conditional score matrix $s(\theta; u)$, which depends on u. Therefore, in order to bound the integral of the influence function, we will bound $\int_0^\infty s(\theta; u)G(du)$. Notice also that each term in the score matrix contains the quantity $(y+1)P_{y+1}/P_y$, so we focus on bounds for these quantities in the following three lemmas.

Lemma 1 *For* $y = 0$,

$$(y+1)\frac{P_{y+1}}{P_y} = \frac{P_1}{P_0} = \mu(1+2\tau\mu)^{-1/2}.$$

Proof As presented in Seshadri (1993), the probability generating function for the PIG(μ, $1/\tau$) model is

$$\sum_{y=0}^{\infty} P_y z^y = \exp\left(\tau^{-1}[1-(1-2\tau\mu(z-1))^{1/2}]\right).$$

From this function, we find that the following recursive relationship exists: $P_1 = \mu(1+2\tau\mu)^{-1/2} P_0$. Thus,

$$(y+1)\frac{P_{y+1}}{P_y} = \frac{P_1}{P_0} = \mu(1+2\tau\mu)^{-1/2}.$$

□

We refer the reader to the Appendix for proofs of the next two lemmas. The first lemma gives an upper bound for the ratio of Poisson inverse Gaussian probabilities; the second, a lower bound.

Lemma 2 *Let* $\nu = +\sqrt{\tau^{-1}(\tau^{-1}+2\mu)}$. *Then for* $y > 0$,

$$(y+1)\frac{P_{y+1}}{P_y} \leq (1+2\tau)^{-1/2}(1+\nu^{-1})(2y-1).$$

Lemma 3 *Let* $\nu = +\sqrt{\tau^{-1}(\tau^{-1}+2\mu)}$. *Then for* $y > 0$,

$$(y+1)\frac{P_{y+1}}{P_y} \geq (1+2\tau)^{-1/2}\left[1+\left(\frac{1}{1+2\nu}\right)^{y}\right].$$

Using the above two lemmas, we now state our main results. The first result gives uniform bounds for the integral of the conditional expected score function for β_0.

Proposition 3 *Let* $\mu = \mu(X)$ *and* $\nu = +\sqrt{\tau^{-1}(\tau^{-1}+2\mu)}$. *The integral of the conditional expected score function for* β_0*, given by*

$$\int_0^{\infty} \mathbb{E}_X \mathbb{E}_{Y|X,U}\left\{ Y - (Y+1)\frac{P_{Y+1}}{P_Y} \,\middle|\, X, U = u \right\} G(du),$$

is uniformly bounded above by

$$\mathbb{E}_X\left\{\mu - (1+2\tau)^{-1/2}\left[1+\mathcal{L}_G\left(\frac{2\nu\mu}{1+2\nu}\right)\right]\right\}$$

and below by

$$\mathbb{E}_X[\mu - (1+2\tau)^{-1/2}(1+\nu^{-1})(2\mu - 1)],$$

where $\mathcal{L}_G(\cdot)$ denotes the Laplace transform of G.

Proof Using Lemma 3,

$$\begin{aligned}
&\mathbb{E}_X\ \mathbb{E}_{Y|X,U}\left\{Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right\} \\
&\quad \le \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{\left[Y - (1+2\tau)^{-1/2}\left(1 + \left(\frac{1}{1+2\nu}\right)^Y\right)\right]\middle| X, U = u\right\} \\
&\quad = \mathbb{E}_X\left[u\mu - (1+2\tau)^{-1/2}\left(1 + \exp\left(-\frac{2\nu\mu u}{1+2\nu}\right)\right)\right],
\end{aligned}$$

where the second term comes from evaluating the probability generating function of a Poisson random variable. Therefore, we have the following:

$$\begin{aligned}
&\int_0^\infty \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right\} G(du) \\
&\quad \le \int_0^\infty \mathbb{E}_X\left[u\mu - (1+2\tau)^{-1/2}\left(1 + \exp\left(-\frac{2\nu\mu u}{1+2\nu}\right)\right)\right] G(du) \\
&\quad = \mathbb{E}_X\int_0^\infty \left[u\mu - (1+2\tau)^{-1/2}\left(1 + \exp\left(-\frac{2\nu u\mu}{1+2\nu}\right)\right)\right] G(du) \\
&\quad = \mathbb{E}_X\left\{\mu - (1+2\tau)^{-1/2}\left[1 + \mathcal{L}_G\left(\frac{2\nu\mu}{1+2\nu}\right)\right]\right\}.
\end{aligned}$$

This gives our upper bound. For a lower bound, recall from Lemma 2 that we have

$$\begin{aligned}
&\mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right\} \\
&\quad \ge \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{\left[Y - (1+2\tau)^{-1/2}\left(1 + \frac{1}{\nu}\right)(2Y - 1)\right]\middle| X, U = u\right\} \\
&\quad = \mathbb{E}_X[u\mu - (1+2\tau)^{-1/2}(1+\nu^{-1})(2u\mu - 1)].
\end{aligned}$$

Integrating with respect to G, we have that

$$\begin{aligned}
&\int_0^\infty \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{Y - (Y+1)\frac{P_{Y+1}}{P_Y}\middle| X, U = u\right\} G(du) \\
&\quad \ge \int_0^\infty \mathbb{E}_X[u\mu - (1+2\tau)^{-1/2}(1+\nu^{-1})(2u\mu - 1)]G(du) \\
&\quad = \mathbb{E}_X\int_0^\infty [u\mu - (1+2\tau)^{-1/2}(1+\nu^{-1})(2u\mu - 1)]G(du)
\end{aligned}$$

$$= \mathbb{E}_X[\mu - (1+2\tau)^{-1/2}(1+\nu^{-1})(2\mu - 1)]$$

as claimed. □

The next result for β_1 is analogous to the previous one. Its proof directly follows that of Proposition 3 after including the covariate X.

Proposition 4 *Let $\mu = \mu(X)$ and $\nu = +\sqrt{\tau^{-1}(\tau^{-1} + 2\mu)}$. The integral of the conditional expected score function for β_1, given by*

$$\int_0^\infty \mathbb{E}_X \mathbb{E}_{Y|X,U} \left\{ X \left[Y - (Y+1)\frac{P_{Y+1}}{P_Y} \right] \middle| X, U = u \right\} G(du),$$

is uniformly bounded above by

$$\mathbb{E}_X \left\{ X\mu - X(1+2\tau)^{-1/2} \left[1 + \mathcal{L}_G\left(\frac{2\nu\mu}{1+2\nu} \right) \right] \right\}$$

and below by

$$\mathbb{E}_X[X\mu - X(1+2\tau)^{-1/2}(1+\nu^{-1})(2\mu - 1)],$$

where $\mathcal{L}_G(\cdot)$ denotes the Laplace transform of G.

The last main result concerns bounds for the expected conditional score function for τ.

Proposition 5 *Let $\mu = \mu(X)$ and $\nu = +\sqrt{\tau^{-1}(\tau^{-1} + 2\mu)}$. The integral of the conditional expected score function for τ, given by*

$$\tau^{-2} \int_0^\infty \mathbb{E}_X \mathbb{E}_{Y|X,U} \left\{ (1+\tau\mu) \left(\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y} - 1 \right) \middle| X, U = u \right\} G(du),$$

is uniformly bounded above by

$$\tau^{-2}\mathbb{E}_X \left\{ (1+\tau\mu)[\mu^{-1}(1+2\tau)^{-1/2}(1+\nu^{-1})(2\mu - 1) - 1] \right\}$$

and below by

$$\tau^{-2}\mathbb{E}_X \left\{ (1+\tau\mu) \left(\mu^{-1}(1+2\tau)^{-1/2} \left[1 + \mathcal{L}_G\left(\frac{2\nu\mu}{1+2\nu} \right) \right] - 1 \right) \right\},$$

where $\mathcal{L}_G(\cdot)$ denotes the Laplace transform of G.

Proof By Lemma 2, we have

$$\begin{aligned}
&\tau^{-2}\mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left(\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y}-1\right)\middle| X, U=u\right\}\\
&\leq \tau^{-2}\mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left[\mu^{-1}(1+2\tau)^{-1/2}\left(1+\nu^{-1}\right)(2Y-1)-1\right]\middle| X, U=u\right\}\\
&= \tau^{-2}\mathbb{E}_X(1+\tau\mu)\left[\mu^{-1}(1+2\tau)^{-1/2}\left(1+\nu^{-1}\right)(2u\mu-1)-1\right].
\end{aligned}$$

Integrating with respect to G, we have that

$$\begin{aligned}
&\tau^{-2}\int_0^\infty \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left(\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y}-1\right)\middle| X, U=u\right\}\\
&\qquad\leq \tau^{-2}\int_0^\infty \mathbb{E}_X(1+\tau\mu)\left[\mu^{-1}(1+2\tau)^{-1/2}\left(1+\nu^{-1}\right)(2\mu u-1)-1\right]dG(u)\\
&\qquad= \tau^{-2}\mathbb{E}_X\int_0^\infty (1+\tau\mu)\left[\mu^{-1}(1+2\tau)^{-1/2}\left(1+\nu^{-1}\right)(2\mu u-1)-1\right]dG(u)\\
&\qquad= \tau^{-2}\mathbb{E}_X\left\{(1+\tau\mu)[\mu^{-1}(1+2\tau)^{-1/2}\left(1+\nu^{-1}\right)(2\mu-1)-1]\right\}.
\end{aligned}$$

This gives our upper bound. For a lower bound, we use Lemma 3 to obtain

$$\begin{aligned}
&\tau^{-2}\mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left(\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y}-1\right)\middle| X, U=u\right\}\\
&\quad\geq \tau^{-2}\mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left(\mu^{-1}(1+2\tau)^{-1/2}\left[1+(1+2\nu)^{-Y}\right]-1\right)\middle| X, U=u\right\}\\
&\quad= \tau^{-2}\mathbb{E}_X(1+\tau\mu)\left(\mu^{-1}(1+2\tau)^{-1/2}\left[1+\exp\left(\frac{-2\nu u\mu}{1+2\nu}\right)\right]-1\right),
\end{aligned}$$

where the second term comes from evaluating the probability generating function of a Poisson random variable. Therefore,

$$\begin{aligned}
&\tau^{-2}\int \mathbb{E}_X\mathbb{E}_{Y|X,U}\left\{(1+\tau\mu)\left(\mu^{-1}(Y+1)\frac{P_{Y+1}}{P_Y}-1\right)\middle| X, U=u\right\}G(du)\\
&\qquad\geq \tau^{-2}\int_0^\infty \mathbb{E}_X(1+\tau\mu)\left(\mu^{-1}(1+2\tau)^{-1/2}\left[1+\exp\left(\frac{-2\nu u\mu}{1+2\nu}\right)\right]-1\right)dG(u)\\
&\qquad\geq \tau^{-2}\mathbb{E}_X\int_0^\infty (1+\tau\mu)\left(\mu^{-1}(1+2\tau)^{-1/2}\left[1+\exp\left(\frac{-2\nu u\mu}{1+2\nu}\right)\right]-1\right)dG(u)\\
&\qquad= \tau^{-2}\mathbb{E}_X\left\{(1+\tau\mu)\left(\mu^{-1}(1+2\tau)^{-1/2}\left[1+\mathcal{L}_G\left(\frac{2\nu\mu}{1+2\nu}\right)\right]-1\right)\right\}.
\end{aligned}$$

□

The above three propositions have shown that the integral of the conditional expected score matrix $s(\theta; u)$ is uniformly bounded. These bounds were found by using properties of Bessel functions and by relating the Bessel functions to the Poisson inverse Gaussian probabilities.

Note that the random effect U is positive (with probability 1); therefore, the Laplace transform of G in Propositions 3–5 results in the uniformity of the respective bounds. Hence, for the PIG$(1, 1/\tau)$ model, $\hat{\theta}$ is robust to mixing distribution misspecification.

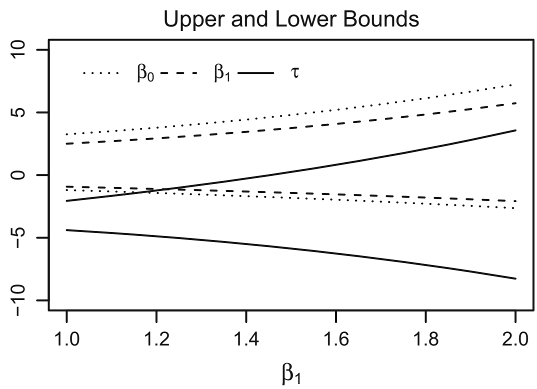

Fig. 1 Upper and lower bounds for $\int s(\theta; u)G(du)$, the integral of the conditional expected score function, where $\theta = (\beta_0, \beta_1, \tau)$. Shown are the bounds corresponding to β_0 (dotted line), β_1 (dashed line), and τ (solid line) when $\beta_0 = 0.5$, $\tau = 1$, $\beta_1 = (1, 1.1, \ldots, 2)$, and $G \sim \Gamma(1, 1)$

3.3 Discussion of upper and lower bounds

The upper and lower bounds presented in Propositions 3–5 involve $\mathcal{L}_G(\cdot)$, the Laplace transform of the contaminating distribution G, which is equivalent to the moment generating function (mgf) of G if it exists. We also note that the bounds involve expectations of $\mu(X) = \exp(\beta_0 + \beta_1 X)$ with respect to the distribution of X, requiring the mgf of X as well.

To illustrate the bounds presented in Propositions 3–5, we assume that $G \sim \Gamma(1, 1)$, $X \sim$ Uniform[0.5, 1], $\beta_0 = 0.5$, $\tau = 1$, and $\beta_1 = (1, 1.1, \ldots, 2)$. The upper and lower bounds for the integrals of the conditional expected score function of (12) for β_0, β_1, and τ are plotted against β_1 in Fig. 1. Note that an increase in β_1 corresponds to an increase in $\mu(X)$. We use the `integrate` function in *R* for numerical integration (R Core Team 2017). For smaller values of β_1, the upper and lower bounds are closer, and they grow farther apart as β_1 increases. In particular, note the similarities between the upper and lower bounds for β_0 and β_1. The distance between the upper and lower bounds for τ increases at a faster rate than the distance between the bounds corresponding to the regression parameters. This example suggests that $\hat{\tau}$ may be more sensitive to misspecification of the mixing distribution.

4 Simulations

In this simulation study, we explore how well the asymptotic results of the previous section describe the true performance of MLEs when the mixing distribution is misspecified in various ways. We simulate Poisson mixtures and compute MLEs of the parameters under an assumed inverse Gaussian mixing distribution. These estimates are computed using the *R* package `gamlss` (Stasinopoulos et al. 2017), which produces MLEs for a Poisson inverse Gaussian regression model.

As noted by Hilbe (2014), the Poisson gamma (or negative binomial) regression model is the most commonly used model for overdispersed count data. Therefore, for these simulations, we let $F \sim \mathrm{IG}(1, 1/\tau)$ denote the nominal inverse Gaussian mixing distribution and $G \sim \Gamma(1/\tau, \tau)$ denote the contaminating gamma distribution. MLEs are computed for an assumed PIG$(1, 1/\tau)$ model.

We begin by generating 1000 *iid* covariates $\{X_i\} \sim \mathrm{N}(0, 1)$, that are fixed throughout the simulation, and by forming the regression structure $\beta_0 + X_i\beta_1$. Next, we generate 1000 random effects $\{U_i\}$, from either the inverse Gaussian distribution or a misspecified mixing distribution, such that $\mathbb{E}(U_i) = 1$ and $\mathrm{Var}(U_i) = \tau$. Then, 1000 sample responses $\{Y_i\}$ are generated, conditionally on U_i and X_i so that

$$\mathbb{E}(Y_i|X_i, U_i) = U_i \exp\{\beta_0 + X_i\beta_1\}.$$

We simulate 1000 replicated samples with the same X_i and estimate β_0, β_1, and τ via maximum likelihood, and we examine the statistical behavior of these estimates using a Monte Carlo approach.

Table 1 shows Monte Carlo estimates for $\beta_0 = 0.5$, $\beta_1 = 1$, and $\tau = (0.25, 0.5, 1, 2)$ when there is a small amount of contamination ($\epsilon = 0.01$) and complete misspecification ($\epsilon = 1$) of the mixing distribution. Although the theory presented in Sect. 3.2 is for infinitesimal perturbations of the mixing distribution, we consider complete misspecification of the mixing distribution which may be done in practice.

When $\epsilon = 0.01$, the bias of $\hat{\beta}_0$ and $\hat{\beta}_1$ is very small, regardless of the true τ. The bias of $\hat{\tau}$ is also small (-0.0096, -0.0141,-0.0374, -0.0907, respectively), yet it increases with τ. The standard errors of the regression estimates increase gradually as τ increases. In contrast, the standard error of $\hat{\tau}$ increases rapidly as τ increases. These results suggest that $\hat{\tau}$ is less robust than $\hat{\beta}_0$ and $\hat{\beta}_1$. However, these results are not surprising since τ is the variance of the mixing distribution which has been contaminated.

When $\epsilon = 1$, the bias of $\hat{\beta}_0$ and $\hat{\beta}_1$ is still small, and their standard errors increase gradually. Note that the standard errors of the regression estimates under complete misspecification are similar to those when $\epsilon = 0.01$. However, the bias of $\hat{\tau}$ is substantial; its standard error grows rapidly as τ increases, so that for $\tau \geq 1$, the bias is the dominant component of the mean squared error of $\hat{\tau}$.

Figure 2 gives normal probability plots of the MLEs for $\theta = (\beta_0, \beta_1, \tau) = (0.5, 1, 2)$ when the mixing distribution is completely contaminated. In general, the plots suggest normality of the estimates, though there is some deviation from a straight line in the tails of the distribution for $\hat{\tau}$. In Fig. 3, the histograms of estimates for $\tau = (0.25, 0.5, 1, 2)$ when $\beta_0 = 0.5$ and $\beta_1 = 1$ also suggest normality of the MLEs; however, the plots reveal the considerable bias of $\hat{\tau}$ for $\tau \geq 1$.

5 Summary

We have focused on the Poisson inverse Gaussian model and have examined the effects of mixing distribution misspecification on MLEs. Extending the influence function approach of Hampel (1974) to the unobservable random effects, we computed bounds for the Gâteaux derivatives of regression parameter estimators and the random effects variance estimator. We showed that the MLEs are robust against small perturbations

Table 1 True and estimated parameter values under assumed inverse Gaussian mixing distribution, $\epsilon = .01$ and $\epsilon = 1$

Parameter	β_0	β_1 $\epsilon = 0.01$	τ	β_0	β_1 $\epsilon = 1$	τ
True	0.5	1	0.25	0.5	1	0.25
Estimated	0.4987	1.0003	0.2404	0.5021	0.9966	0.2609
SE ($\times 10^2$)	(3.25)	(2.92)	(3.47)	(3.26)	(3.12)	(3.76)
True	0.5	1	0.5	0.5	1	0.5
Estimated	0.4993	0.9998	0.4859	0.4982	0.9970	0.5651
SE ($\times 10^2$)	(3.65)	(3.51)	(5.71)	(3.77)	(3.72)	(6.59)
True	0.5	1	1	0.5	1	1
Estimated	0.4980	1.0022	0.9626	0.5005	0.9811	1.2965
SE ($\times 10^2$)	(4.26)	(4.30)	(10.39)	(4.67)	(4.46)	(14.05)
True	0.5	1	2	0.5	1	2
Estimated	0.4953	1.0003	1.9093	0.4985	0.9424	3.1598
SE ($\times 10^2$)	(5.29)	(5.26)	(20.87)	(6.38)	(5.76)	(36.91)

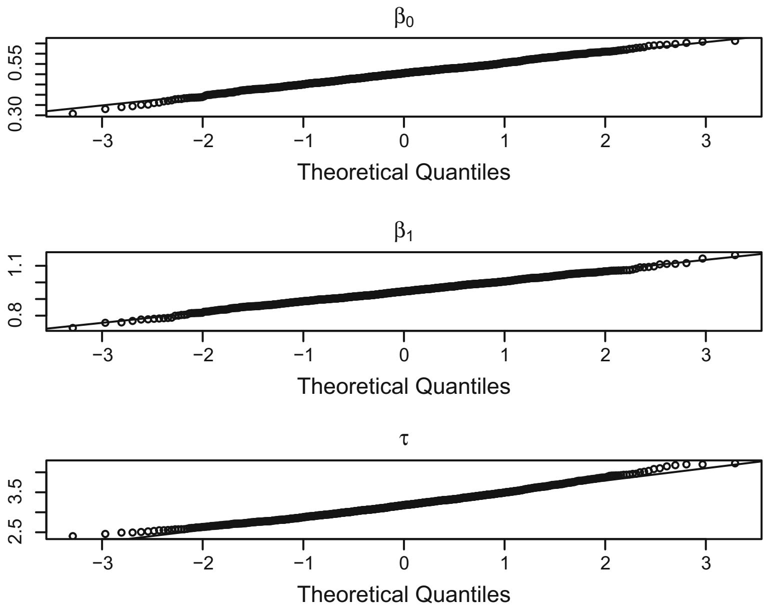

Fig. 2 Normal probability plots of 1000 simulated MLEs for Poisson-gamma data under assumed inverse Gaussian mixing distribution with $\beta_0 = 0.5$, $\beta_1 = 1$, and $\tau = 2$

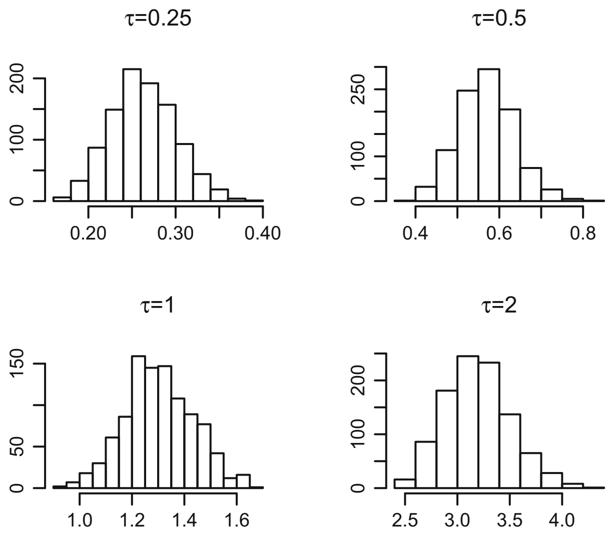

Fig. 3 Histograms of 1000 simulated MLEs for the random effects variance τ for Poisson-gamma data under assumed inverse Gaussian mixing distribution with $\beta_0 = 0.5$, $\beta_1 = 1$, and $\tau = (0.25, 0.5, 1, 2)$

of the mixing distribution, provided that the first two moments exist. Using properties of Bessel functions and exploiting the recursive nature of ratios of Poisson inverse Gaussian probabilities, we have computed bounds for the Gâteaux derivatives of the MLEs.

The simulation study of Sect. 4 considered the practical case of complete misspecification of the mixing distribution. For the regression parameter estimates, our study supports conclusions of McCulloch and Neuhaus (2013) and others who argue that the choice of the mixing distribution has little impact. However, for the MLE of the variance of the random effects distribution, our simulation results suggest that mixing distribution misspecification can have a substantial impact as claimed by Litière et al. (2008). Therefore, when analyzing over-dispersed count data, one can be reasonably confident when using maximum likelihood to estimate regression parameters; however, if one believes that modeling assumptions may be incorrect, alternatives to maximum likelihood estimation should be considered, particularly when estimating the random effects variance.

Acknowledgements The authors thank the Editor, reviewers, Dr. Dennis Boos (North Carolina State University) and Dr. Kimberly F. Sellers (Georgetown University) for helpful comments and suggestions that significantly improved this paper.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix: Upper and lower bounds for Poisson inverse Gaussian probability ratios

Below we provide proofs to Lemmas 2 and 3, in which we find an upper bound and a lower bound for Poisson inverse Gaussian probability ratios.

Lemma 2. Let $\nu = +\sqrt{\tau^{-1}(\tau^{-1} + 2\mu)}$. Then for $y > 0$,

$$(y+1)\frac{P_{y+1}}{P_y} \leq (1+2\tau)^{-1/2}(1+\nu^{-1})(2y-1).$$

Proof Equation (7) gives us the following relationship:

$$\begin{aligned}(y+1)\frac{P_{y+1}}{P_y} &= (y+1)\left\{\frac{(1+2\tau)^{-1/2}}{y+1}\frac{K_{y+\frac{1}{2}}(\nu)}{K_{y-\frac{1}{2}}(\nu)}\right\}\\ &= (1+2\tau)^{-1/2}\frac{K_{y+\frac{1}{2}}(\nu)}{K_{y-\frac{1}{2}}(\nu)}. \end{aligned} \tag{13}$$

From Abramowitz and Stegun (1972), we have that for $k = 0, \pm 1, \pm 2, \ldots$

$$K_{k+\frac{1}{2}}(\nu) = C(\nu)\sum_{r=0}^{k}\frac{(k+r)!}{r!(k-r)!(2\nu)^r}, \tag{14}$$

where $C(\nu) = \sqrt{\pi/(2\nu)}\exp(-\nu)$. Therefore, we may write

$$\begin{aligned}\frac{K_{y+\frac{1}{2}}(\nu)}{C(\nu)} &= \sum_{r=0}^{y}\frac{(y+r)!}{r!(y-r)!(2\nu)^r}\\ &= \sum_{r=0}^{y}\frac{(y-1+r)!}{r!(y-1-r)!(2\nu)^r}\frac{(y+r)}{(y-r)}\\ &= \frac{(2y)!}{y!(2\nu)^y} + \sum_{r=0}^{y-1}\frac{(y-1+r)!}{r!(y-1-r)!(2\nu)^r}\frac{(y+r)}{(y-r)}\\ &\leq \frac{(2y)!}{y!(2z)^y} + (2y-1)\sum_{r=0}^{y-1}\frac{(y-1+r)!}{r!(y-1-r)!(2\nu)^r}\\ &\leq \frac{(2y)!}{y!(2\nu)^y} + \frac{(2y-1)K_{y-\frac{1}{2}}(\nu)}{C(\nu)}.\end{aligned}$$

Substituting in the numerator of (13) we find that

$$(y+1)\frac{P_{y+1}}{P_y} \leq (1+2\tau)^{-1/2}\left[\frac{C(\nu)(2y)!/(y!(2\nu)^y) + (2y-1)K_{y-\frac{1}{2}}(\nu)}{K_{y-\frac{1}{2}}(\nu)}\right]$$

$$= (1+2\tau)^{-1/2}\left[\frac{C(\nu)(2y)!}{y!(2\nu)^y K_{y-\frac{1}{2}}(\nu)} + 2y - 1\right]$$
$$\leq (1+2\tau)^{-1/2}(1+\nu^{-1})(2y-1), \qquad (15)$$

where the last line uses the following inequality:

$$\frac{y!(2\nu)^y K_{y-\frac{1}{2}}(\nu)}{C(\nu)(2y)!} = \frac{y!(2\nu)^y}{(2y)!}\sum_{r=0}^{y-1}\frac{(y-1+r)!}{r!(y-1-r)!(2\nu)^r}$$
$$\geq \frac{y!(2\nu)^y}{(2y)!}\frac{(2y-2)!}{(y-1)!(2\nu)^{y-1}}$$
$$= \frac{\nu}{2y-1}.$$

□

Lemma 3 Let $\nu = +\sqrt{\tau^{-1}(\tau^{-1}+2\mu)}$. Then for $y > 0$,

$$(y+1)\frac{P_{y+1}}{P_y} \geq (1+2\tau)^{-1/2}\left[1 + \left(\frac{1}{1+2\nu}\right)^y\right].$$

Proof Recall from Eq. (13) that

$$(y+1)\frac{P_{y+1}}{P_y} = (1+2\tau)^{-1/2}\frac{K_{y+\frac{1}{2}}(\nu)}{K_{y-\frac{1}{2}}(\nu)}. \qquad (16)$$

Using (14) with $C(\nu) = \sqrt{\pi/(2\nu)}\exp(-\nu)$, we may write

$$\frac{K_{y-\frac{1}{2}}(z)}{C(\nu)} = \sum_{r=0}^{y-1}\frac{(y-1+r)!}{r!(y-1-r)!(2\nu)^r}$$
$$= \sum_{r=0}^{y-1}\frac{(y+r)!}{r!(y-r)!(2\nu)^r}\frac{(y-r)}{(y+r)}$$
$$\leq \sum_{r=0}^{y-1}\frac{(y+r)!}{r!(y-r)!(2\nu)^r}$$
$$= \sum_{r=0}^{y}\frac{(y+r)!}{r!(y-r)!(2z)^r} - \frac{(2y)!}{y!(2\nu)^y}$$
$$= \frac{K_{y+\frac{1}{2}}(\nu)}{C(\nu)} - \frac{(2y)!}{y!(2\nu)^y}.$$

Therefore, substituting into (16), we have the following lower bound:

$$\begin{aligned}(y+1)\frac{P_{y+1}}{P_y} &\geq (1+2\tau)^{-1/2}\frac{K_{y+\frac{1}{2}}(\nu)}{K_{y+\frac{1}{2}}(\nu)-C(\nu)(2y)!/(y!(2\nu)^y)}\\ &= (1+2\tau)^{-1/2}\frac{1}{1-C(\nu)(2y)!/(y!(2\nu)^y K_{y+\frac{1}{2}}(\nu))}\\ &\geq (1+2\tau)^{-1/2}\left[1+\frac{C(\nu)(2y)!}{y!(2\nu)^y K_{y+\frac{1}{2}}(\nu)}\right], \qquad (17)\end{aligned}$$

where we use the elementary inequality $1/(1-x) \geq 1+x$. Notice that

$$\begin{aligned}\frac{y!(2\nu)^y K_{y+\frac{1}{2}}(\nu)}{C(\nu)(2y)!} &= \sum_{r=0}^{y}\frac{(y+r)!}{r!(y-r)!(2\nu)^r}\frac{y!(2\nu)^y}{(2y)!}\\ &= \sum_{r=0}^{y}\frac{(y+r)!}{(2y)!}\frac{y!(2\nu)^{y-r}}{r!(y-r)!}\\ &= \sum_{r=0}^{y}\frac{(y+r)!}{(2y)!}\binom{y}{r}(2\nu)^{y-r}\\ &\leq \sum_{r=0}^{y}\binom{y}{r}(2\nu)^{y-r}\\ &= (1+2\nu)^y.\end{aligned}$$

Therefore, (17) becomes

$$(y+1)\frac{P_{y+1}}{P_y} \geq (1+2\tau)^{-1/2}\left[1+\left(\frac{1}{1+2\nu}\right)^y\right]$$

by substitution. □

References

Abramowitz M, Stegun I (eds) (1972) Handbook of mathematical functions. Dover, New York
Bickel PJ (1984) Robust regression based on infinitesimal neighbourhoods. Ann Stat 12:1349–1368
Bickel PJ, Doksum KA (2001) Mathematical statistics: basic ideas and selected topics. Prentice Hall, Upper Saddle River
de Bruijn NG (1953) The difference-differential equation $f'(x) = e^{\alpha}x + \beta f(x-1)$, i, ii. Indag Math 15(449–458):459–464
Dean CB, Nielsen JD (2007) Generalized linear mixed models: a review and some extensions. Lifetime Data Anal 13(4):497–512
Dean C, Lawless JF, Willmot GE (1989) A mixed Poisson-inverse-Gaussian regression model. Canad J Stat 17:171–181
Feller W (1966) An introduction to probability theory and its applications, vol 2. Wiley, New York
Gustafson P (1996) The effect of mixing-distribution misspecification in conjugate mixture models. Canad J Stat 24:307–318

Hampel FR (1974) The influence curve and its role in robust estimation. J Am Stat Assoc 69:383–393
Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust statistics: the approach based on influence functions. Wiley, New York
Heagerty PJ, Zeger SL (2000) Marginalized multilevel models and likelihood inference (with comments and a rejoinder by the authors). Stat Sci 15(1):1–26
Heckman J, Singer B (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econom J Econom Soc 52(2):271–320
Hilbe JM (2014) Modeling count data. Cambridge University Press, New York
Hougaard P, Lee ML, Whitmore GA (1997) Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. Biometrics 53:1225–1238
Huber PJ (1981) Robust statistics. Wiley, New York
Karlis D, Xekalaki E (2005) Mixed Poisson distributions. Int Stat Rev 73(1):35–58
Litière S, Alonso A, Molenberghs G (2008) The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models. Stat Med 27(16):3125–3144
McCulloch CE, Neuhaus JM (2013) Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. Biostatistics 14:477–490
McCulloch CE, Searle SR, Neuhaus JM (2008) Generalized, linear and mixed models. Wiley, New York
Neuhaus JM, Hauck WW, Kalbfleisch JD (1992) The effects of mixture distribution misspecification when fitting mixed-effects logistic models. Biometrika 79:755–762
Ong SH (1998) A note on the mixed Poisson formulation of the Poisson-inverse Gaussian distribution. Commun Stat Simul Comput 27(1):67–78
R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, 2017. https://www.R-project.org
Rieder H (1994) Robust Asymptot Stat. Springer, New York
Seshadri V (1993) The inverse Gaussian distribution: a case study in exponential families. Clarendon, Oxford
Seshadri V (1999) The inverse Gaussian distribution: statistical theory and applications. Springer, New York
Shaban SA (1981) Computation of the Poisson-inverse Gaussian distribution. Commun Stat Theory Methods 10(14):1389–1399
Shoukri MM, Asyali MH, VanDorp R, Kelton DF (2004) The Poisson inverse Gaussian regression model in the analysis of clustered counts data. J Data Sci 2:17–32
Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, De Bastiani F (2017) Flexible regression and smoothing: using GAMLSS in R. Chapman and Hall/CRC, London
Teugels JL, Willmot G (1987) Approximations for stop-loss premiums. Insur Math Econ 6:195–202
Van de Ven R, Weber NC (1995) Log-linear models for mean and dispersion in mixed Poisson regression models. Aust J Stat 37(2):205–216
Verbeke G, Molenberghs G (2013) The gradient function as an exploratory goodness-of-fit assessment of the random-effects distribution in mixed models. Biostatistics 14:477–490
Weems KS, Smith PJ (2004) On robustness of maximum likelihood estimates for Poisson-lognormal models. Stat Probab Lett 66:189–196
Willmot GE (1990) Asymptotic tail behaviour of Poisson mixtures with applications. Adv Appl Probab 22:147–159
Zha L, Lord D, Zou Y (2016) The Poisson inverse Gaussian (PIG) generalized linear regression model for analyzing motor vehicle crash data. J Transp Saf Secur 8(1):18–35